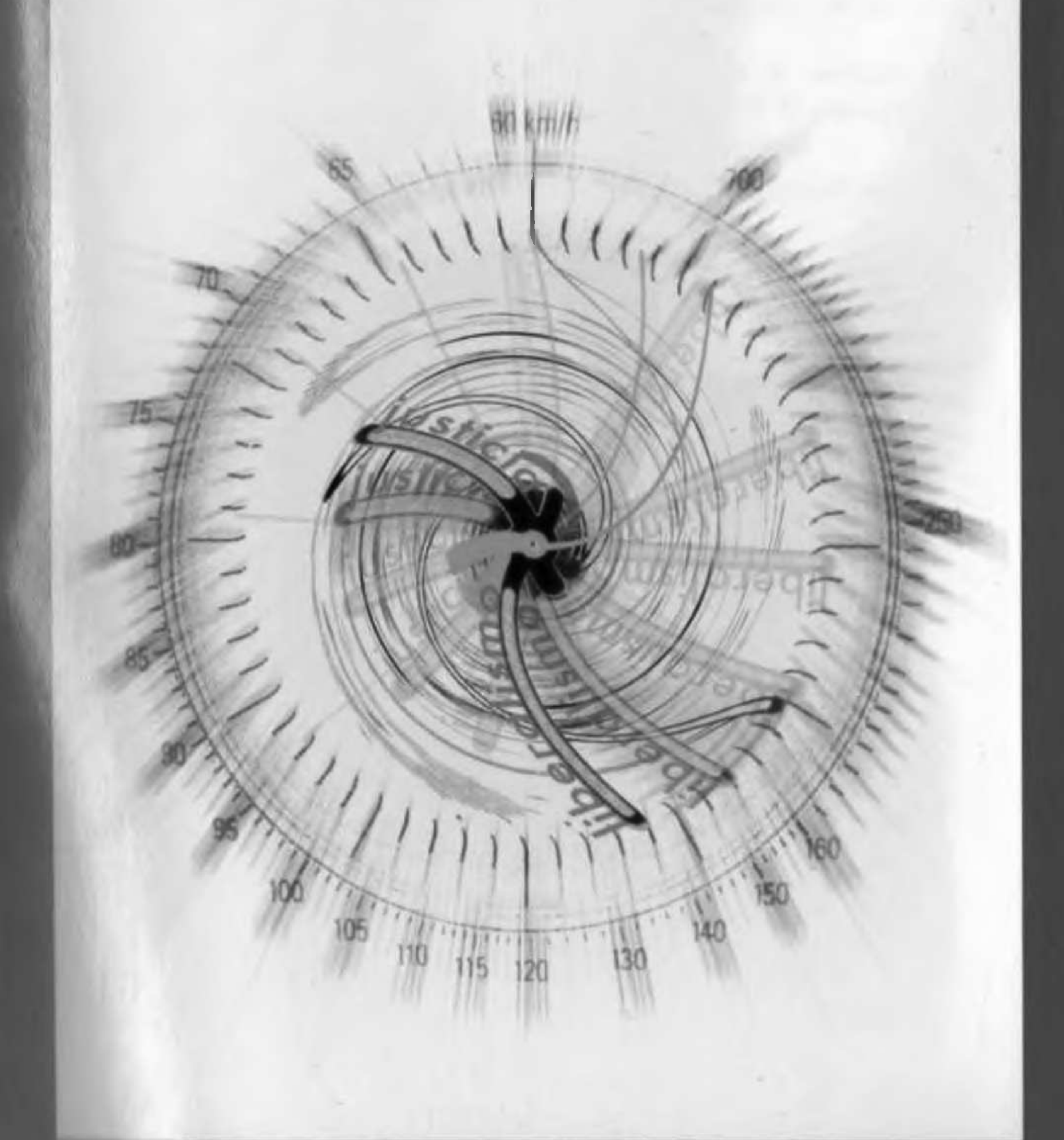
Filosofía del Derecho

Michael Sandel

El liberalismo y los límites de la justicia



gedisa

Filosofía del Derecho
Serie CLA·DE·MA



Michael Sandel
El liberalismo y los límites de la justicia

Una sociedad liberal no pretende imponer una única forma de vida, sino permitir a sus ciudadanos elegir de la manera más libre posible sus propios valores y propósitos. Por ello debe gobernarse con principios de justicia que no presuponen ninguna visión particular de lo que es una buena vida. Sin embargo, ¿pueden encontrarse tales principios? Y si no es posible, ¿cuáles son las consecuencias para la justicia como ideal moral y político?

Éstas son las preguntas que Michael Sandel plantea en su aguda y penetrante crítica al liberalismo contemporáneo. Después de situar los horizontes del liberalismo moderno en la tradición kantiana de la filosofía del derecho, analiza la expresión actual más influyente de aquélla que es la teoría elaborada por John Rawls. En la discusión de sus aspectos más importantes, Sandel explora los límites del liberalismo que subyacen en su concepción de la persona y aboga por una comprensión más profunda de la comunidad que la admitida por el pensamiento liberal.

Esta segunda edición de la obra incluye nuevos comentarios sobre la importante discusión de los años ochenta y noventa entre liberalismo y comunitarismo y un análisis de la evolución posterior del pensamiento de John Rawls.

«Se trata de una nueva y auténtica voz filosófica ... La elegante argumentación del libro de Sandel describe lo que yo entiendo como la realidad de la experiencia moral.»
Michael Walzer, *«The New Republic»*

«La notable obra de Sandel ... descubre las tensiones y contradicciones en una gran parte del liberalismo contemporáneo.... Una filosofía política a la altura de cómo debería escribirse, que enfrenta nuestras creencias morales con nuestra mejor comprensión de la naturaleza humana.»
Charles Taylor, McGill University

Michael Sandel es profesor adjunto de Teoría Política de la Universidad de Harvard, autor de *Democracy's Discontent* (1996), compilador de *Liberalism and its Critics* (1984) y publicó numerosos ensayos sobre teoría política y del derecho.


ISBN 84-7432-706-7
9 788474 327069

Cla•De•Ma/Derecho

Colección dirigida por
Ernesto Garzón Valdés y Jorge Malem Seña

En la actualidad, la filosofía del derecho ya no es un discurso hermético y reservado a los especialistas de cada una de sus parcelas. Además, las modernas herramientas de análisis no sólo abren nuevos accesos metodológicos a los juristas mismos, sino que inciden también en cuestiones que afectan las responsabilidades acerca de nuestra civilización, como son las transgresiones de los derechos humanos o la desproporción entre el crecimiento de la riqueza y su necesaria distribución. La colección pretende contribuir al debate que se está desarrollando en la Filosofía del derecho en el ámbito de la lengua castellana con la publicación de obras y recopilaciones de ensayos de autores internacionales cuyos planteamientos merecen una mayor difusión en nuestra área cultural. La intención es ofrecer un panorama lo más amplio posible y dar preferencia a propuestas abiertas a la discusión teórica más que presentar soluciones definitivas.

A. AARNIO, E. GARZÓN VALDÉS, J. UUSITALO *compiladores*	*La normatividad del derecho*
CARLOS SANTIAGO NINO	*La constitución de la democracia deliberativa*
RODOLFO VÁZQUEZ *compilador*	*Derecho y moral*
OWEN FISS	*La ironía de la libertad de expresión*
DENNIS F. THOMPSON	*La ética política y el ejercicio de cargos públicos*
DAVID LYONS	*Aspectos morales de la teoría jurídica*
J. G. RIDDALL	*Teoría del derecho*
HILARY PUTNAM	*El pragmatismo: un debate abierto*
RICCARDO GUASTINI	*Distinguiendo* Estudios de teoría y metateoría del derecho
J. L. MACKIE	*Ética* La invención de lo bueno y lo malo
BRUCE ACKERMAN	*La política del diálogo liberal*
JEAN PIERRE DUPUY	*El sacrificio y la envidia* El liberalismo frente a la justicia social
DAVID GAUTHIER	*La moral por acuerdo*
BRIAN BARRY	*Teorías de la justicia*
ROBERT A. DAHL	*Después de la revolución* La autoridad en las sociedades democráticas
JON ELSTER	*Justicia local*
RONALD DWORKIN	*El imperio de la justicia*

EL LIBERALISMO Y LOS LÍMITES DE LA JUSTICIA

Michael J. Sandel

Traducción: María Luz Melon



Título original: *Liberalism and the Limits of Justice*

Publicado por el Press Syndicate de la University of Cambridge
Cambridge, Nueva York, Victoria

Traducción: María Luz Melon

Diseño de cubierta: Juan Santana

Primera edición, junio del 2000, Barcelona

Paseo Bonanova, 9 1º-1ª
08022 Barcelona (España)
Tel.: 93 253 09 04
Fax: 93 253 09 05
Correo electrónico: gedisa@gedisa.com
http://www.gedisa.com

ISBN 84-7432-706-7
Depósito legal: B. 28.396/2000

Impreso por: Limpergraf
c/ Mogoda, 29-31 – 08210 Barberà del Vallès

Impreso en España
Printed in Spain



A mis padres

Índice

Agradecimientos

Esta obra comenzó a gestarse en Oxford a finales de los años setenta, una época estimulante para el estudio de la filosofía política, especialmente en el Balliol College. Tengo una deuda de gratitud para con los profesores y amigos de aquellos años, quienes me enseñaron mucho de lo que he escrito aquí. Quiero expresar mi agradecimiento en primer lugar a William Connolly, Richard Fallon, Donald Herzog, Steven Lukes, David Miller, Alan Montefiore, Judith Shklar y Charles Taylor, quienes leyeron una versión anterior de este ensayo y me hicieron valiosos comentarios y críticas. Sus consejos fueron siempre beneficiosos, y si aún subsistiera algún punto débil, no es por su culpa. Algunas personas, sin embargo, se implicaron de un modo más profundo. Entre mis colegas en el extranjero, quiero mencionar a Ronald Beiner, Richard Fallon, y Scott Matheson, porque han asumido una responsabilidad especial. Las conversaciones y viajes que realicé con ellos me ayudaron a dar forma a mis pensamientos, y me mostraron un compañerismo intelectual del más alto nivel. Sin embargo, mi mayor deuda la tengo con aquellos de mis maestros cuya influencia está más presente en esta obra: Ronald Dworkin, cuyos argumentos exigían una mejor respuesta de la que yo podía ofrecer en aquel momento; Charles Taylor, quien amplió los horizontes angloamericanos y nos enseñó la relevancia de Aristóteles y Hegel; y especialmente Alan Montefiore, quien transformó la filosofía en algo ineludible y me inició en este camino hace seis años.

Introducción

El liberalismo y la primacía de la justicia

Este ensayo se ocupa del liberalismo, pero más concretamente de una de sus variantes que tiene mucha relevancia en la filosofía política, moral y jurídica actual. En este liberalismo, las nociones de justicia, de equidad y de los derechos individuales desempeñan un papel central; además debe gran parte de su fundamentación filosófica a Kant. En tanto ética que afirma la prioridad de lo justo sobre lo bueno, y que se define de manera clásica por oposición a las concepciones utilitaristas, la mejor manera de definirlo es como «liberalismo deontológico»; un nombre resonante para una doctrina de la que espero que sea conocida.

El «liberalismo deontológico» es, sobre todo, una teoría acerca de la justicia. En particular, es una teoría que sostiene la primacía de la justicia entre los ideales políticos y morales. Su tesis central podría formularse de la siguiente manera: la sociedad, compuesta por una pluralidad de individuos cada uno de los cuales tiene sus propios fines, intereses y concepciones del bien, está mejor ordenada cuando se gobierna por principios que no presuponen ninguna concepción particular del bien *per se*. Lo que justifica estos principios regulativos por encima de todo no es el hecho de que maximicen el bienestar social ni que promuevan el bien, sino más bien el que estén en conformidad con el concepto de lo *justo*, que es una categoría moral que precede al bien y es independiente de éste.

Ésta es la versión del liberalismo definida por Kant y por gran parte de la filosofía moral y política contemporáneas, y lo que me propongo es refutar este liberalismo. Frente a la tesis de la primacía de la justicia postularé que existen límites de la justicia y que esto implica que también el liberalismo tenga límites. Me refiero a unos límites no prácticos sino conceptuales. No postulo que sea improbable que la justicia, por noble que sea en principio, alguna vez pueda llegar a verificarse completamente en la práctica; lo que afirmo es que los límites residen en el ideal mismo. El problema de una sociedad que se inspira en la promesa liberal no es simplemente que la justicia esté siempre por realizarse, sino que la concepción es deficiente y la aspiración incompleta. Sin embargo, antes de explorar estos límites debemos considerar más claramente en qué consiste el postulado de la primacía de la justicia.

Los fundamentos del liberalismo: Kant frente a Mill

La primacía de la justicia puede comprenderse de dos modos diferentes, aunque relacionados. El primer modo implica un sentido directamente moral. Este sentido moral afirma que la justicia es primaria porque las demandas de justicia pesan más que otros intereses políticos y morales, sin importar cuán perentorios sean estos otros. Dentro de esta perspectiva, la justicia no es meramente un valor entre otros, que ha de ser evaluado y considerado cuando se den las circunstancias, sino que es la virtud social cardinal, la única que debe cumplirse antes de prestar oídos al reclamo de cualquiera de las demás. Si la felicidad del mundo pudiera promoverse sólo a través de medios injustos, entonces no debería prevalecer la felicidad sino la justicia. Y cuando la justicia culmina en ciertos derechos individuales, ni siquiera el bienestar general podría superarlos.

Pero la primacía de la justicia, tomada sólo en su sentido moral, difícilmente sería suficiente para distinguir a este liberalismo de otras versiones bien conocidas. Muchos pensadores liberales han enfatizado la importancia de la justicia y han insistido en la entidad suprema de los derechos individuales. John Stuart Mill llamó a la justicia «la parte más importante

y mucho más inviolablemente obligatoria que ninguna otra de la moral» (1863: 185). Por su parte, Locke sostuvo que los derechos naturales del hombre son demasiado fuertes como para ser superados por ningún otro derecho del Estado (1690). Pero ninguno de ellos era un liberal deontológico en el sentido más profundo que aquí nos interesa. La verdadera ética deontológica no se ocupa solamente de la moral, sino también de la fundamentación de la moral. Se trata no sólo del peso de la ley moral sino también de los modos cómo se deriva aquello que Kant denominaría su «fundamento determinante» (Kant 1788).

Desde el punto de vista plenamente deontológico, la primacía de la justicia describe no sólo una prioridad moral sino también una forma privilegiada de justificación. Lo justo tiene prioridad sobre lo bueno, no sólo porque sus pretensiones lo preceden, sino también en virtud de que sus principios se derivan independientemente. Esto quiere decir que, a diferencia de otros mandatos prácticos, los principios de la justicia se justifican de un modo que no depende de ninguna visión particular de lo bueno. Por el contrario; debido a su independencia, lo justo limita a lo bueno y establece sus fronteras. «El concepto de lo bueno y malo tiene que ser determinado, no antes de la ley moral (para la cual ese concepto parecía deber ser colocado como fundamento), sino sólo (como aquí ocurre) después de la misma y por la misma» (Kant 1788: 134-135).

Desde el punto de vista de la fundamentación moral, la primacía de la justicia significa lo siguiente: la virtud de la ley moral no consiste en el hecho de que promueva algún objetivo o fin que se presume bueno. Por el contrario, la ley moral tiene las siguientes propiedades: es un fin en sí mismo, precede a todos los otros fines, y es regulativa respecto de ellos. Kant distingue este sentido fundacional de primacía (de segundo orden) respecto del sentido moral (de primer orden) de la siguiente manera:

> Por primado entre dos o más cosas ligadas por razón, entiendo yo la ventaja que una tiene de ser el primer fundamento de determinación de la unión con todas las demás. En sentido práctico estricto, significa la ventaja del interés de la una en cuanto a que a este interés (que no puede ser puesto detrás de ningún otro) está subordinado el interés de las otras (Kant 1788: 174).

Esta oposición también podría presentarse en términos de dos sentidos diferentes de deontología. En su sentido moral, la deontología se opone al *consecuencialismo*, al describir una ética de primer orden que contiene ciertos deberes y prohibiciones categóricos que tienen prioridad incondicional sobre otras cuestiones morales y prácticas. En su sentido fundacional, la deontología se opone a la *teleología* porque describe una forma de justificación según la cual los primeros principios se derivan de un modo que no presupone ningún propósito o fin humano final, ni una concepción determinada de la bondad humana.

Sin lugar a dudas, la primera de las dos ramas de la ética deontológica es la más conocida. Muchos liberales, no sólo los deontológicos, han otorgado una importancia especial a la justicia y a los derechos individuales. Esta actitud suscita una cuestión relativa a la manera en que se relacionan los dos aspectos de la deontología. ¿Puede defenderse el primer tipo de liberalismo sin recurrir al segundo? Mill, por ejemplo, pensaba de esa manera, y argumentó a favor de la posibilidad, incluso la necesidad, de separar ambos tipos de liberalismo.

Mill afirma que tener un derecho es «tener algo cuya posesión debe garantizar la sociedad» (1863: 180). Tan fuerte es la obligación para la sociedad, que mi derecho «asume ese carácter absoluto, esa aparente infinidad e inconmensurabilidad respecto de las otras consideraciones, que constituye la diferencia existente entre el sentimiento de lo justo y lo injusto y entre lo que es ordinariamente conveniente y lo perjudicial» (Mill 1863: 181). Pero si se pregunta por qué la sociedad debe aceptar esta obligación, no es «por otra razón que la utilidad general» (Mill 1863: 180). La justicia se considera oportunamente como «*la parte más importante y mucho más inviolablemente obligatoria que ninguna otra de la moral*», no en razón de una justicia abstracta, sino simplemente porque los requerimientos de la justicia «ocupan un rango más elevado en la escala de la utilidad social y, por tanto, poseen una obligatoriedad superior a la de las otras» (Mill 1863: 188).

> Es importante establecer que yo hago proceder cualquier ventaja que podría derivarse de mi argumento de la idea de una justicia abstracta, como una cosa independiente de la utilidad. Considero a la utilidad como la última instancia de apelación para todas las cuestiones éticas; pero debe ser la utilidad en el sentido

más extendido, fundada en los intereses permanentes del hombre como un ser progresivo (Mill 1849: 485).

La importancia predominante de la justicia y de los derechos individuales los hace «más absolutos e imperativos» que otras demandas. Pero aquello que los hace importantes en primer lugar es el servicio que prestan a la utilidad social, su fundamento último. «Toda acción se realiza con vistas a un fin, y parece natural suponer que las reglas de una acción deban tomar su carácter y color del fin al cual se subordinan» (Mill 1863: 134). En la perspectiva utilitarista, los principios de la justicia, como todos los demás principios morales, obtienen su carácter y su matiz a partir de la felicidad entendida como fin, ya que «preguntarse por los fines es preguntarse qué cosas son deseables» y dado que la felicidad es deseable, es de hecho «la única cosa deseable como fin», porque «la gente la desea de hecho» (Mill 1863: 164). Aquí surgen con claridad la fundamentación teleológica y los presupuestos psicológicos del liberalismo postulado por Mill.

Para Kant, por el contrario, los dos aspectos de la deontología están estrechamente vinculados. En su ética y su metafísica argumenta rigurosamente contra la posibilidad de establecer un aspecto sin que al mismo tiempo se establezca el otro. El punto de vista kantiano sugiere al menos dos objeciones precisas contra una posición como la de Mill (y la de los actuales «utilitaristas de la regla»). La primera afirma que los fundamentos utilitaristas no son confiables. La segunda, que los fundamentos que no son confiables, cuando se trata de la justicia, pueden resultar coercitivos e injustos.

No se puede confiar en el utilitarismo porque ningún fundamento meramente empírico, utilitario o de otro tipo, puede asegurar absolutamente la primacía de la justicia y la integridad de los derechos individuales. Un principio que debe presuponer ciertos deseos e inclinaciones no puede estar menos condicionado que esos mismos deseos. Pero tanto nuestros deseos como los medios de satisfacerlos varían de manera característica entre distintas personas, al igual que con respecto a una misma persona con el transcurso del tiempo. En consecuencia, cualquier principio que dependa de esos deseos debe ser igualmente contingente. Por lo tanto, «todos los principios prácticos que suponen un objeto (materia) de la facultad de desear como fundamento de determinación de la voluntad, son

todos ellos empíricos y no pueden proporcionar ley práctica alguna» (Kant 1788: 104). Si la utilidad es el fundamento determinante —aun la «utilidad en el sentido más amplio»—, entonces en principio debe haber casos en los que el bienestar general anule a la justicia en lugar de asegurarla.

Mill, de hecho, acepta este punto, pero se preguntaría si la justicia necesariamente debería ser tan incondicionalmente privilegiada. Reconoce que la consideración utilitaria no establece la prioridad de la justicia de manera absoluta, y que pueden surgir casos particulares «en los que algún otro deber social sea tan importante como para predominar sobre cualquiera de las máximas generales de la justicia» (Mill 1863: 188-189). Pero si a través de esta limitación se promueve la felicidad de la humanidad, ¿qué fundamentos podría haber para afirmar la primacía de la justicia de manera más absoluta?[1]

La respuesta de Kant sería que debemos rechazar incluso las excepciones que se hacen en nombre de la felicidad de la humanidad, puesto que el no afirmar de manera absoluta la primacía de la justicia conduce a la injusticia y a la coerción. Aun si el deseo de felicidad fuera universalmente compartido, no podría servir como fundamento para la ley moral. Las personas seguirían teniendo concepciones diferentes sobre la naturaleza de la felicidad, y el establecer cualquier concepción particular como regla impondría a algunos la concepción de otros, y negaría al menos a algunos la libertad de promover sus propias concepciones. Se crearía una sociedad en la cual algunos individuos serían forzados por los valores de otros, en lugar de una en la cual las necesidades de cada individuo se armonizasen con los fines de todos. «Los hombres tienen diferentes puntos de vista sobre qué es y en qué consiste el fin empírico de la felicidad, de modo que, en lo que respecta a la felicidad, su voluntad no puede englobarse bajo ningún principio común ni por lo tanto bajo ninguna ley externa que armonice con la libertad de todos» (Kant 1793: 73-4).

1. Mill afirma que la justicia simplemente «es» cualquier cosa que la utilidad requiera. En los casos en que las máximas generales de la justicia son superadas, «solemos decir, no que la justicia debe ceder el paso a algún otro principio moral, sino que lo que es justo en los casos ordinarios, no lo es en un caso particular por razón de ese otro principio. Por medio de esta formulación útil del lenguaje, se salvaguarda el carácter de inviolabilidad atribuido a la justicia, y nos libramos de la necesidad de sostener que puede haber injusticias laudables» (Mill 1863: 189).

Para Kant, la prioridad de lo justo se «deriva enteramente del concepto de libertad en la mutua relación externa de los seres humanos, y no tiene nada que ver con el fin que se considera natural para todos los hombres (v.g. el propósito de lograr la felicidad) o con los medios reconocidos de lograr este fin» (Kant 1793: 73). La justicia debe tener un fundamento anterior a todos los fines empíricos. Ni siquiera una unión fundada sobre algún fin común que todos los miembros de la sociedad compartieran sería adecuada. Sólo una unión «como un fin en sí mismo que todos ellos deben compartir y que es de ese modo un deber absoluto y primario en todas las relaciones externas, cualesquiera sean, entre seres humanos» puede asegurar la justicia y evitar la coerción ejercida sobre algunos individuos por las convicciones de los otros. Sólo en una unión semejante nadie puede «obligarme a ser feliz de acuerdo con su concepción del bienestar de los otros» (Kant 1793: 73-4). Sólo cuando soy gobernado por principios que no presuponen ningún fin particular, soy libre para perseguir mis propios fines, que son consistentes con una libertad semejante para todos los otros individuos.

En la perspectiva kantiana, las dos ramas de la ética deontológica están unidas. La prioridad moral de la justicia se hace posible (y necesaria) por su prioridad fundacional. La justicia tiene primacía sobre los demás valores porque sus principios se derivan independientemente. La ley moral no se ve implicada en la promoción de distintos intereses y fines contingentes, como los otros principios prácticos. La ley moral no presupone ninguna concepción particular de lo bueno. En virtud de que el fundamento de la justicia se da antes que todos los fines meramente empíricos, ésta no pierde su privilegio frente a lo bueno y, de esta manera, establece sus límites.

Ahora bien, una solución de este tipo plantea la cuestión de cuál podría llegar a ser realmente un fundamento para la justicia. Si debe ser un fundamento anterior a todos los propósitos y fines, no condicionado aún por «las propiedades particulares de la naturaleza humana» (Kant 1785: 42), ¿dónde sería posible encontrar un fundamento tal? Si tenemos en cuenta las condiciones estrictas de la ética deontológica, se sugiere que la ley moral prácticamente no se fundamentaría en ninguna otra cosa, ya que cualquier requisito previo material debilitaría su prioridad. «¡Deber!» pregunta Kant en su tono más lírico, «¿cuál es el origen digno de ti? ¿Dónde se halla la raíz de

tu noble ascendencia, que rechaza orgullosamente todo parentesco con las inclinaciones?» (Kant 1788: 151).

Su respuesta es que el fundamento de la ley moral debe encontrarse en el sujeto de la razón práctica y no en su objeto. Esto es, en un sujeto capaz de tener una voluntad autónoma. Por lo tanto, no es un fin empírico, «el sujeto de los fines, esto es, el ser racional mismo, no debe nunca ponerse por fundamento de las acciones como simple medio, sino como suprema condición limitativa en el uso de todos los medios...» (Kant 1785: 50). Esto significa que sólo «el sujeto de todos los fines posibles» puede producir lo justo, ya que es también el sujeto de una voluntad autónoma. Sólo una concepción tal podría ser «algo que eleva al hombre sobre sí mismo como una parte del mundo sensible» y lo capacita para participar en un ideal. Es decir que lo capacita para participar en un ámbito no condicionado, totalmente independiente de nuestras inclinaciones sociales y psicológicas. Y sólo esta independencia absoluta nos puede proporcionar el distanciamiento que necesitamos para poder elegir libremente por nosotros mismos, sin vernos coartados por las circunstancias contingentes. En la perspectiva deontológica, lo más importante no son los fines que elegimos sino nuestra capacidad para elegirlos. Y esta capacidad, al ser anterior a cualquier fin particular que pueda anhelarse, reside en el sujeto. «No es ninguna otra más que la personalidad, es decir, la libertad e independencia del mecanismo de toda la naturaleza, considerada esa libertad, sin embargo, al mismo tiempo como una facultad de un ser que está sometido a leyes puras prácticas peculiares, es decir, dadas por su propia razón» (Kant 1788: 151).

El concepto de un sujeto que es anterior a sus objetos e independiente de ellos ofrece un fundamento para la ley moral. Ésta no necesita teleología o psicología alguna, a diferencia de los fundamentos meramente empíricos. De este modo se completa poderosamente la visión deontológica. Como lo justo es anterior a lo bueno, de ese modo el sujeto es anterior a sus fines. Para Kant, estas prioridades paralelas explican «de una vez el fundamento que ha ocasionado todos los errores de los filósofos con respecto al principio supremo de la moral. Pues ellos buscaban un objeto de la voluntad para hacer de él la materia y el fundamento de una ley.» Esta elección estaba sin embargo obligada a dejar a sus primeros principios confundidos en la heteronomía. «Ellos, en cambio, hubiesen debido buscar

primeramente una ley que determinase a priori e inmediatamente la voluntad, y sólo después, según esa ley, el objeto» (Kant 1788: 135). Si hubiesen seguido ese camino, habrían llegado a la distinción entre el sujeto y el objeto de la razón práctica. De ese modo, habrían obtenido un fundamento para la justicia que fuera independiente de cualquier objeto particular.

Si el argumento a favor de la primacía de la justicia ha de tener éxito, si lo justo tiene que ser anterior a lo bueno en los sentidos morales y fundacionales entrelazados que hemos distinguido, entonces también debe funcionar adecuadamente alguna versión del postulado en favor de la primacía del sujeto. Hasta aquí esto resulta claro. Queda por demostrar la posibilidad de defender este último postulado. ¿Cómo sabemos que existe un sujeto que puede identificarse antes que los objetos que anhela y aparte de ellos? Esta pregunta cobra un particular interés una vez que recordamos que el argumento a favor de la prioridad del sujeto no es un argumento empírico. Si fuera un argumento empírico, difícilmente podría llevar a cabo la tarea que lo recomienda en primer lugar para la ética deontológica.

El sujeto trascendental

Kant ofrece dos argumentos en apoyo de su noción de sujeto. El primero es epistemológico, el segundo es un argumento práctico. Ambos constituyen formas de argumento «trascendentales» porque actúan a través de la búsqueda de los presupuestos subyacentes a ciertos aspectos aparentemente indispensables de nuestra experiencia. El argumento epistemológico investiga los presupuestos del autoconocimiento. Comienza con la idea de que un sujeto no puede conocer todo lo que es posible conocer sobre sí mismo meramente por medio de la observación o de la introspección. Esto se debe a que cuando se lleva a cabo la introspección, todo lo que es posible observar son los aportes de los propios sentidos; que es posible conocerse sólo en tanto objeto de experiencia, como el portador de éste o aquél deseo, inclinación, querer, disposición, etc. Pero este tipo de autoconocimiento está destinado a la limita-

ción, puesto que nunca puede capacitar al sujeto para llegar más allá del flujo de las apariencias y descubrir *de qué* lo son. «E incluso no le es lícito al hombre pretender conocerse a sí mismo, tal como es en sí, por el conocimiento que de sí tiene mediante la sensación interna» (Kant 1785: 58). La introspección, o «sensación interna», por sí misma nunca podría proporcionar conocimiento sobre algo que esté detrás de estas apariencias, ya que cualquier aporte se disolvería instantáneamente en una nueva apariencia. Sin embargo, debemos presuponer algo más. «[El hombre] tiene que admitir sobre esa constitución de su propio sujeto, compuesta de meros fenómenos, alguna otra cosa que esté a su base [...] esto es, su "yo" tal como sea en sí [...]» (Kant 1785: 59).

Este algo más, que no podemos conocer empíricamente sino que debemos al menos presuponer como la condición para conocer cualquier cosa, es el sujeto mismo. El sujeto es ese algo «subyacente», que antecede a cualquier experiencia particular, que unifica nuestras diversas percepciones y las coloca reunidas en una conciencia singular. Proporciona el principio de unidad sin el cual las percepciones de sí que tiene el sujeto serían nada más que un flujo de representaciones desconectadas y siempre cambiantes, las percepciones de nadie. En tanto que no podemos aprehender empíricamente este principio, debemos presuponer su validez si queremos lograr que nuestro autoconocimiento tenga algún sentido.

> La idea de que todas las representaciones dadas en la intuición me pertenecen, es en consecuencia equivalente a la idea de que las unifico en una autoconciencia, o al menos que de esa manera las puedo unificar; y aun cuando esta idea no es ella misma la conciencia de la *síntesis* de las representaciones, sin embargo presupone la posibilidad de esa síntesis. En otras palabras, sólo en tanto puedo aprehender la multiplicidad de las representaciones en una conciencia, puedo llamarlas a todas ellas *mías*. Ya que de otra manera debería tener un sí mismo tan diverso y polifacético como las representaciones de las cuales soy consciente. (Kant 1787: 154)

El descubrimiento de que debo comprenderme como un sujeto además de como un objeto de experiencia sugiere dos modos diferentes de concebir las leyes que gobiernan mis acciones. De ese modo conduce a partir del argumento epistemológico hacia un argumento ulterior, de tipo práctico, a favor de la prioridad

del sujeto. En tanto *objeto* de experiencia, pertenezco al mundo sensible; mis acciones, al igual que los movimientos de todos los otros objetos, se determinan por las leyes de la naturaleza y por las regularidades de causa y efecto. En tanto *sujeto* de experiencia, por el contrario, habito un mundo inteligible o suprasensible; en éste, independiente de las leyes de la naturaleza, soy capaz de tener autonomía, capaz de actuar de acuerdo con una ley que me doy a mí mismo.

Sólo a partir de este segundo punto de vista puedo considerarme libre, «pues la independencia de las causas determinantes del mundo sensible (independencia que la razón tiene siempre que atribuirse) es libertad» (Kant 1785: 59). Si yo fuese un ser completamente empírico, no sería capaz de tener libertad, ya que todo ejercicio de la voluntad estaría condicionado por el deseo de algún objeto. Toda elección sería heterónoma, regida por la prosecución de algún fin. Mi voluntad nunca podría ser una causa primera, sino meramente el efecto de alguna causa anterior, el instrumento de uno u otro impulso o inclinación. En tanto nos consideremos libres a nosotros mismos, no podemos considerarnos como seres meramente empíricos. «Pues ahora ya vemos que, cuando nos pensamos como libres, nos incluimos en el mundo inteligible, como miembros de él, y conocemos la autonomía de la voluntad...» (Kant 1785: 60). Y así la noción de un sujeto anterior a la experiencia e independiente de ésta, tal como lo requiere la ética deontológica, aparece no sólo como posible sino como indispensable; una presuposición necesaria para la posibilidad de autoconocimiento y de libertad.

Ahora podemos ver más claramente en qué consiste el postulado a favor de la primacía de la justicia desde el punto de vista de la ética deontológica. Desde el punto de vista kantiano la primacía de lo justo es al mismo tiempo moral y fundacional. Se fundamenta en el concepto de un sujeto dado antes que sus fines, un concepto que se sostiene como indispensable para nuestra comprensión de nosotros mismos en tanto seres autónomos, que eligen libremente. La sociedad se ordena mejor cuando es gobernada por principios que no presuponen ninguna concepción particular de lo bueno, ya que cualquier otro ordenamiento no lograría respetar a las personas en tanto seres capaces de elección. Los consideraría más bien como objetos antes que como sujetos, como medios antes que como fines en sí mismos.

Los temas deontológicos encuentran una expresión similar en gran parte del pensamiento liberal contemporáneo. Así, «...los derechos asegurados por la justicia no están sujetos a regateos políticos ni al cálculo de intereses sociales» (Rawls, 1971: 20), sino que «funcionan como cartas de triunfo que se reservan los individuos» (Dworkin 1978: 136) frente a las políticas que impondrían alguna visión particular de lo bueno sobre la sociedad en su conjunto. «Puesto que los ciudadanos de una sociedad difieren en sus concepciones», el gobierno no cumple con la obligación de respetarlos como iguales «si prefiere una concepción a otra, ya sea porque los funcionarios crean que una es intrínsecamente superior, o porque dicha concepción es la que sostiene el grupo más poderoso o más numeroso» (Dworkin 1978: 127). En comparación con lo bueno, los conceptos de lo justo y lo injusto «tienen una naturaleza independiente y dominante porque establecen nuestra posición básica como entidades que eligen libremente». Más importante que cualquier elección, el valor de ser persona «es el presupuesto y el substrato del concepto de elección. Y por eso es que las normas relativas al respeto de la persona no pueden ser comprometidas, porque estas normas son absolutas con respecto a los diversos fines que elegimos perseguir» (Fried 1978: 8-9, 29).

Este liberalismo —al menos en sus versiones contemporáneas— se presenta, en virtud de su independencia respecto de los presupuestos teleológicos y psicológicos comunes, como inmune a la mayoría de las controversias a las cuales han sido tradicionalmente vulnerables las teorías políticas, especialmente con respecto a cuestiones de la naturaleza humana y del significado de la vida en el bien. De ese modo se exige que «el liberalismo no se apoye en ninguna teoría especial de la personalidad» (Dworkin 1978: 142), que sus presupuestos claves no involucren «ninguna teoría particular de la motivación humana» (Rawls 1971: 154), que «los liberales, en tanto tales, sean indiferentes» a los modos de vida que los individuos elijan (Dworkin 1978: 143), y que, para aceptar el liberalismo, uno «no necesite tomar una posición sobre una serie de *grandes temas* de carácter altamente controvertido» (Ackerman 1980: 361).

Pero si bien ciertos «grandes temas» de la filosofía y de la psicología se encuentran más allá del interés del liberalismo deontológico, esto se da sólo en la medida en que este libera-

lismo establece su controversia desde otro lugar. Como hemos visto, evita confiar en una teoría de la persona en particular, al menos en el sentido tradicional de atribuir a todos los seres humanos una determinada naturaleza, o ciertos deseos e inclinaciones esenciales, tales como la sociabilidad o el egoísmo. Sin embargo existe otro sentido en el que este liberalismo implica una cierta teoría de la persona: su centro de interés lo constituye, no el objeto de los deseos humanos, sino el sujeto del deseo y la manera en que se constituye este sujeto.

Para que la justicia sea primaria, deben ser verdaderas determinadas cosas con respecto a nosotros. Debemos ser criaturas de un cierto tipo, relacionadas con la circunstancia humana de una cierta manera. En particular, debemos asumir nuestra circunstancia siempre desde una cierta distancia, seguramente condicionados, pero parte de nosotros siempre antecederá a cualquier condición. Solamente de esta manera podemos considerarnos como sujetos y también como objetos de experiencia, como agentes y no sólo como instrumentos de los propósitos que perseguimos. El liberalismo deontológico supone que podemos, e incluso que debemos, comprendernos como independientes en este sentido. Mi planteamiento es que esto no es posible para nosotros, y que es en la parcialidad de esta imagen de nosotros mismos donde se pueden encontrar los límites de la justicia.

Pero entonces, ¿en qué se equivoca la teoría deontológica de la persona? ¿Cómo es que sus errores debilitan paulatinamente la primacía de la justicia? ¿Qué virtud rival aparece en el momento en que se encuentran los límites de la justicia? Éstas son las preguntas que esta obra aspira a responder. A fin de presentar el escenario para mi argumento, primero será de gran utilidad considerar otras dos críticas que pueden hacerse a la perspectiva kantiana.

La objeción sociológica

La primera podría denominarse la objeción sociológica, ya que comienza por destacar la influencia omnipresente de las condiciones sociales en la formulación de los valores individuales y las medidas de política. Alega que el liberalismo se equi-

voca, ya que la neutralidad es imposible, y la neutralidad es imposible porque, sin importar cuánto lo intentemos, nunca podremos escapar totalmente de los efectos de nuestro condicionamiento. Todos los órdenes políticos encarnan en este sentido *algunos* valores; el problema es *de quién* son los valores que prevalecen, y quién gana y quién pierde como resultado. La proclamada independencia del sujeto deontológico es una ilusión liberal. Interpreta mal la naturaleza fundamentalmente «social» del hombre, el hecho de que somos seres condicionados «de principio a fin». No hay ningún punto de excepción, ni ningún sujeto trascendental capaz de ubicarse fuera de la sociedad o fuera de la experiencia. En todo momento somos lo que hemos llegado a ser, una concatenación de deseos e inclinaciones que no deja nada libre para habitar un dominio noumenal. La prioridad del sujeto puede sólo significar la prioridad del individuo, sesgando de ese modo la concepción en favor de los valores individualistas, conocidos para la tradición liberal. La justicia sólo *aparece* como primaria porque este individualismo produce típicamente postulados contradictorios. Por consiguiente los límites de la justicia consistirían en la posibilidad de cultivar aquellas virtudes cooperativas, tales como el altruismo y la benevolencia, que logran que el conflicto sea menos perentorio. Pero éstas son precisamente las virtudes que resulta menos probable que surjan en una sociedad fundada en presupuestos individualistas. En resumen, la falsa promesa del liberalismo consiste en el ideal de una sociedad gobernada por principios neutrales. Afirma los valores individualistas mientras pretende una neutralidad que nunca es posible lograr.

Pero la objeción sociológica no logra apreciar la fuerza del punto de vista deontológico en muchos aspectos. Primero, no comprende la neutralidad que el liberalismo alega ofrecer. Lo que es neutral con respecto a los principios de lo justo no es el que se admitan todos los valores y fines posibles, sino que se deriven de forma tal que no dependan de ningún valor ni fin particular. Sin duda, una vez que se encuentran disponibles los principios de la justicia de esa manera derivados, ciertos fines se excluyen (difícilmente serían *regulativos* si no fueran incompatibles con *nada*), pero sólo aquéllos que sean injustos, es decir, sólo aquellos fines que resulten inconsistentes con principios que no dependan para su propia validez de la validez de un modo de vida determinado. Su neutralidad describe su fundamento, no su efecto.

Pero aún su efecto es de varias importantes maneras menos restrictivo que lo que sugiere la objeción sociológica. El altruismo y la benevolencia, por ejemplo, son totalmente compatibles con este liberalismo, y no hay nada en sus postulados que desaconseje cultivarlos. La prioridad del sujeto no declara que estemos gobernados por el egoísmo; sólo que cualquier interés que tengamos debe ser el interés de algún sujeto. Desde el punto de vista de lo justo, yo soy libre para buscar mi propio bien o el bien de los otros, mientras que no actúe injustamente. Y esta restricción no tiene que ver con el egoísmo o con el altruismo sino con el interés dominante en asegurar una libertad similar para otros. Las virtudes cooperativas no son de ninguna manera inconsistentes con este liberalismo.

Finalmente, no resulta clara la manera en que la objeción sociológica se propone refutar la noción deontológica de independencia. Si esto significa ofrecer una objeción psicológica, entonces no puede alcanzar el punto de vista deontológico, que efectúa una aserción epistemológica. La independencia del sujeto no significa que yo puedo, como un hecho psicológico, convocar en cualquier momento la indiferencia requerida para superar mis prejuicios o superar mis convicciones, sino que mis valores y mis fines no definen mi identidad; que debo considerarme como el portador de un «yo» distinto de mis valores y mis fines, cualesquiera sean éstos.

Si, por el otro lado, la objeción sociológica significa una crítica a esta exigencia epistemológica, no resulta claro cuál podría ser su base. Quizás Hume fue el que más se acercó a describir un «yo» completamente condicionado empíricamente, tal como la objeción sociológica requiere, cuando describe al «yo» como «un grupo o conjunto de percepciones diferentes, que se suceden con una inconcebible rapidez, y se encuentran en permanente flujo y movimiento» (Hume 1739: 252). Pero como Kant argumentaría posteriormente, «ningún "yo" fijo y constante puede presentarse en este flujo de apariencias internas». Para dar sentido a la continuidad del «yo» a lo largo del tiempo, debemos suponer algún principio de unidad que «precede a toda experiencia, y hace posible la experiencia misma» (Kant 1781: 136). De hecho el propio Hume anticipó esta dificultad cuando admitió que no podía finalmente justificar esos principios «que unen nuestras sucesivas percepciones en nuestro pensamiento o consciente» (Hume 1739: 636). No obstante lo problemático

que pueda parecer el sujeto trascendental kantiano, la objeción sociológica no parece estar lo suficientemente preparada como para ofrecer una crítica efectiva. La epistemología que debe presuponer difícilmente será más plausible.

Deontología con rostro humeano

La segunda postura crítica formula una dificultad más profunda relacionada con el sujeto kantiano. Como la primera, surge a partir de una visión empirista. Pero a diferencia de ésta, procura reforzar al liberalismo deontológico en lugar de oponérsele. De hecho esta segunda postura no es tanto una objeción al punto de vista kantiano como una reformulación afín. Acepta la prioridad de lo justo sobre lo bueno, e incluso afirma la prioridad del «yo» sobre sus fines. Esta visión se aparta de Kant en su negativa a considerar que un «yo» lógicamente anterior e independiente de sus fines sólo pueda ser trascendental, o sujeto noumenal, careciendo completamente de fundamento empírico. Esta deontología «revisionista» aprehende el espíritu de gran parte del liberalismo contemporáneo, y encuentra su expresión más completa en el trabajo de John Rawls. «Para desarrollar una concepción kantiana de la justicia que sea viable», escribe, «la fuerza y contenido de la doctrina de Kant debe separarse de su fundamento en el idealismo trascendental» y reconstruirse dentro de los «cánones de un empirismo razonable» (Rawls 1977: 165).

Para Rawls, la concepción kantiana adolece de arbitrariedad y falta de claridad, ya que resulta confusa la manera en que un sujeto abstracto, no corporeizado, podría sin arbitrariedad producir determinados principios de justicia, o cómo en cualquier caso la legislación de un sujeto como éste se aplicaría a los seres humanos efectivos en el mundo fenoménico. La metafísica idealista, a pesar de toda su ventaja moral y política, cede demasiado a lo trascendente, y al formular un dominio noumenal reafirma la primacía de la justicia sólo al precio de negarle su condición humana.

De esa manera Rawls tomó como proyecto preservar las enseñanzas deontológicas kantianas a través del reemplazo de la oscura prosa germana por una metafísica domesticada, menos

posible de ser calificada de arbitraria y análoga al temperamento angloamericano. Su propuesta consiste en derivar los primeros principios a partir de una situación hipotética de elección (la «posición original»), caracterizada por condiciones tales que produzcan un resultado determinado, adecuado a seres humanos reales. No prevalece el reino de los fines sino las circunstancias ordinarias de la justicia, que toma prestadas de Hume. No es un futuro moral que cada vez se aleja más, sino un presente firmemente instalado en la circunstancia humana, el que da ocasión a la justicia. Si el resultado ha de ser la deontología, será una deontología con un rostro humeano.[2]

> La teoría de la justicia trata a su vez de presentar un proceso natural, traducción de la concepción kantiana del reino de los fines y de las nociones de autonomía e imperativo categórico. De este modo, la estructura subyacente de la doctrina de Kant se separa de sus bases metafísicas, de manera que pueda verse más claramente y se presente libre de objeciones. (301)[3]

La cuestión de si la metafísica de Kant es un «entorno» separable o un presupuesto ineludible de las aspiraciones políticas y morales que Kant y Rawls comparten (v.g., si este último es capaz de concentrarse en la política liberal sin la confusión metafísica) es uno de los debates centrales que plantea la concepción de Rawls. Este ensayo argumenta que el propósito de Rawls no alcanza el éxito, y que el liberalismo deontológico no puede rescatarse de las dificultades asociadas con el sujeto kantiano. La deontología con rostro humeano, o bien fracasa como deontología, o bien recrea en la posición original el sujeto no corporeizado que intenta evitar. La justicia no puede ser primaria en el sentido deontológico, porque no nos podemos considerar de manera coherente como el tipo de seres que la ética deontológica (ya sea kantiana o rawlsiana) exige que seamos. Pero la atención hacia este liberalismo no se debe sólo a un interés crítico. El intento de Rawls de situar al «yo» deontológico, reconstruido adecuadamente, nos conduce más allá de la deontología hacia una concepción de comunidad que marca los límites de la justicia y establece lo incompleto del ideal liberal.

2. Mi gratitud a Mark Hulbert por sugerirme esta frase.
3. Todos los números de página que aparecen solos entre paréntesis se refieren a Rawls 1971, *Teoría de la Justicia,* FCE, Buenos Aires, 1993.

I

La justicia y el sujeto moral

Como Kant, Rawls es un liberal deontológico. Su libro adopta la tesis principal de la ética deontológica como postulado central. El hecho de que este postulado no haya sido debatido de manera directa en la voluminosa literatura crítica dedicada a la *Teoría de la Justicia* da cuenta de la firmeza del lugar que ocupa entre los presupuestos morales y políticos de nuestro tiempo. No trata de los principios de lo justo sino de la naturaleza de la justicia misma. La afirmación que abre y cierra la obra, la convicción central que Rawls intenta defender por encima de todo es el postulado de que «la justicia es la primera virtud de las instituciones sociales». Es la consideración singular más importante en la evaluación de la estructura básica de la sociedad y la de la orientación principal del cambio social.

> La justicia es la primera virtud de las instituciones sociales, como la verdad lo es de los sistemas de pensamiento. Una teoría, por muy elegante y económica que fuera, tiene que ser rechazada o revisada si no es verdadera; de igual modo, no importa que las leyes e instituciones estén ordenadas y sean eficientes: si son injustas han de ser reformadas o abolidas. [...] Siendo las primeras virtudes de la actividad humana, la verdad y la justicia no pueden estar sujetas a transacciones. (19-20)
>
> He tratado de exponer una teoría que nos permita comprender y valorar estos sentimientos acerca de la primacía de la justicia. El resultado es la justicia como imparcialidad: articula estas opiniones y respalda su tendencia general. (647)

Lo que me propongo examinar es este postulado en favor de la primacía de la justicia.

La primacía de la justicia y la prioridad del yo

Actualmente la primacía de la justicia es un postulado fuerte, y existe el peligro de que la familiaridad de la idea nos impida ver su audacia. Para comprender por qué esta idea resulta intuitivamente cautivante pero al mismo tiempo profundamente compleja y problemática, podríamos considerar la siguiente reconstrucción del postulado con el fin de captar al mismo tiempo su familiaridad y su fuerza: la justicia no es meramente un valor importante entre otros, que pueda ponderarse y considerarse según lo requiera la ocasión, sino que constituye el *medio* para evaluar y sopesar los valores. Es en este sentido el «valor de valores»,[4] por así decirlo, y no está sujeto al mismo tipo de compensaciones que los valores que regula. La justicia es el criterio por el cual se reconcilian los valores en conflicto y se acomodan las concepciones contradictorias del bien, aunque no siempre se resuelvan. Como tal, debe poseer una cierta prioridad con respecto a aquellos valores y bienes. Ninguna concepción del bien podría prevalecer sobre las exigencias de la justicia, ya que éstas son de un orden cualitativamente diferente; su validez se establece de un modo diferente. Con respecto a los valores sociales, generalmente la justicia permanece separada y distante, tal como un proceso de decisiones justo mantiene la distancia respecto de las pretensiones de quienes a él recurren.

Pero ¿en qué sentido exactamente la justicia, en tanto árbitro de los valores, «debe» ser lógicamente anterior a ellos? Uno de los sentidos de esta prioridad es un «debe» *moral* que surge a partir de la crítica de Rawls a la ética utilitarista. Desde este punto de vista, la prioridad de la justicia es un requisito de la pluralidad esencial de la humanidad y la integridad de los in-

4. La frase es de Alexander Bickel, quien atribuye a la ley una primacía comparable a la condición acordada aquí para la justicia. «El valor irreductible, aun cuando no el exclusivo, es la idea de ley. La ley es más que sólo otra opinión; no porque encarne todos los valores correctos... sino porque es el valor de los valores. La ley es la principal institución por medio de la cual una sociedad puede afirmar sus valores» (1975: 5).

dividuos que ésta comprende. Sacrificar la justicia en favor del bienestar general es violar lo inviolable, es no respetar la distinción entre las personas.

> Cada persona goza de una inviolabilidad fundada en la justicia que incluso el bienestar de la sociedad como un todo no puede atropellar. Por esta razón la justicia niega que la pérdida de libertad de algunos pueda rectificarse por el hecho de que otros compartan un mayor bien. No permite que frente a los sacrificios impuestos a unos pocos tenga mayor peso una cantidad superior de ventajas disfrutadas por muchos. Por tanto, en una sociedad justa, las libertades de la igualdad de ciudadanía se consideran como definitivamente establecidas; los derechos que la justicia garantiza no están sujetos a negociaciones políticas ni cálculos de intereses sociales. (19-20)

Pero hay otro sentido en el cual la justicia *debe* ser lógicamente anterior a los valores que juzga —anterior en el sentido de su derivación independiente— que tiene que ver con un aspecto problemático de los criterios de juicio en general. Se trata de un requisito epistemológico antes que moral, y éste surge a partir del problema de cómo diferenciar entre lo que es un criterio de evaluación y lo que es la materia misma de la evaluación. Como insiste Rawls, necesitamos un «punto arquimédico» a partir del cual podamos evaluar la estructura básica de la sociedad. El problema consiste en justificar en qué lugar se podría imaginar la ubicación de ese punto. Dos posibilidades parecen ofrecerse pero ambas son igualmente insatisfactorias: si los principios de justicia se derivan a partir de los valores o concepciones de lo bueno vigentes en la sociedad, no hay seguridad de que la perspectiva crítica que proporcionan sea más válida que las concepciones que habrían de regular, puesto que al ser un producto de aquellos valores, la justicia estaría sujeta a las mismas contingencias. Como alternativa se podría pensar en un criterio de alguna manera externo a los intereses y valores que prevalecen en la sociedad. Pero si de esta manera nuestra experiencia quedara del todo descalificada como la fuente de tales principios, entonces esta alternativa parecería basarse en presupuestos apriorísticos cuyas credenciales resultarían igualmente sospechosas, aunque por razones opuestas. La primera sería su arbitrariedad por ser contingente, mientras que la segunda sería su arbitrariedad por carecer de fundamentación. Cuando la justicia se deriva a partir de valores existentes los criterios de evalua-

ción se confunden con los objetos a evaluar y no hay manera segura de separar unos de otros. Cuando la justicia viene dada por principios a priori no hay un modo seguro de conectarlos.

Éstos son por tanto los requisitos difíciles y complejos del punto arquimédico, es decir, encontrar una perspectiva que no esté comprometida por su implicación en el mundo ni disociada y por lo tanto descalificada por su desvinculación. «Necesitamos una concepción que nos permita contemplar nuestros objetivos desde lejos» (40), pero no *demasiado* lejos; la perspectiva deseada «no es una perspectiva desde un cierto lugar, más allá del mundo, ni el punto de vista de un ser trascendente; más bien es una cierta forma de pensar y de sentir que las personas racionales pueden adoptar en el mundo» (648).

Antes de considerar la respuesta de Rawls a esta crítica, vale la pena señalar en qué manera su defensa de la primacía de la justicia se relaciona con diversos postulados paralelos presentes en toda su teoría que, tomados en su conjunto, revelan una estructura argumental característica de la ética deontológica como un todo. La noción más general de la prioridad de lo justo sobre lo bueno se relaciona estrechamente con la idea de la primacía de la justicia. Como ésta, la prioridad de lo justo sobre lo bueno aparece inicialmente como un postulado moral de primer orden, en oposición a la doctrina utilitarista, pero en última instancia también llega a adoptar cierto carácter metaético, particularmente cuando Rawls hace una argumentación más genérica a favor de las teorías éticas deontológicas en oposición a las teorías éticas teleológicas.

En tanto postulado estrictamente moral, la prioridad de lo justo sobre lo bueno denota que los principios de lo justo prevalecen invariablemente frente a las consideraciones del bienestar o de la satisfacción de los deseos —no importa cuan intensos sean—, y obligan de antemano a que se satisfaga toda la variedad de deseos y valores que sean realmente legítimos.

> Los principios de lo justo, y por lo tanto de la justicia, ponen un límite al valor de las satisfacciones razonables del bien propio. [...] Esto podemos expresarlo diciendo que, en la justicia como imparcialidad, el concepto de lo justo es previo al de lo bueno. [...] La prioridad de la justicia se explica, en parte, sosteniendo que los intereses que exigen la violación de la justicia carecen de valor. Al no tener méritos en el primer término, no pueden superar sus pretensiones. (49-50)

En la idea de justicia como imparcialidad, a diferencia del utilitarismo, se afirma incondicionalmente el derecho individual a una igualdad de libertades frente a su correspondiente contrario, las preferencias de la mayoría.

> Las intensas convicciones de la mayoría, si son, en realidad, simples preferencias sin fundamento alguno en los principios de la justicia previamente establecidos, no tienen suficiente peso. La satisfacción de estos sentimientos carece de todo valor que pueda colocarse en el platillo de la balanza contra las pretensiones de una libertad igual. [...] Frente a estos principios, nada importa la intensidad de los sentimientos ni el hecho de que sean compartidos por la mayoría. En la interpretación contractual, por tanto, las bases de la libertad se hallan completamente separadas de las preferencias existentes. (497)

Aun cuando Rawls argumenta primero contra las concepciones utilitaristas, su proyecto general es más ambicioso, ya que la justicia como imparcialidad toma posición no sólo frente al utilitarismo, sino frente a todas las teorías teleológicas en tanto tales. Como una exigencia de segundo orden, metaética, la prioridad de lo justo implica que, de «los dos conceptos de ética», lo justo se deriva independientemente de lo bueno, y no al revés. Esta prioridad fundacional permite que lo justo perdure independientemente de los valores y concepciones de lo bueno que prevalezcan, y convierte la posición de Rawls en deontológica antes que teleológica (43; 48).

Una de las primeras consecuencias de primer orden de la ética deontológica consiste en constituir un fundamento más sólido para la libertad igual de los individuos, en comparación con lo que es posible plantear basándose en los presupuestos teleológicos. Es aquí donde surge más claramente la importancia de la deontología para los intereses liberales que nos resultan conocidos. Cuando lo justo es instrumental para la prosecución de algún fin que se considera como anterior, puede justificarse la negación de la libertad de algunos individuos en nombre de un bien que para otros prevalece. Las libertades de la ciudadanía igualitaria están de esa manera «en peligro cuando se fundan en principios teleológicos. El argumento en su favor descansa en cálculos precarios así como en premisas controvertidas e inciertas» (244). En la perspectiva deontológica, «las libertades equitativas tienen un fundamento diferente». Al no ser ya meramente medios para maximizar la sa-

tisfacción o alcanzar algún fin prevaleciente, «estos derechos son asignados para satisfacer los principios de cooperación que reconocerían los ciudadanos al quedar debidamente representados como personas morales» (245), como un fin en sí mismos.

Pero el fracaso en asegurar los derechos a la libertad igual revela una deficiencia más profunda en la concepción teleológica. En la concepción de Rawls, la teleología confunde la relación de lo justo con lo bueno porque concibe erróneamente la relación del «yo» con sus fines. Esto lleva a Rawls a afirmar una nueva prioridad deontológica. Al contrario que en la teleología, lo que resulta más esencial para nuestra condición de persona no son los fines que elegimos, sino nuestra capacidad de elegirlos. Y esta capacidad se localiza en un «yo» que debe ser lógicamente anterior a los fines que elige.

> Y esto sugiere que la estructura de las doctrinas teleológicas es radicalmente equívoca: ya desde el principio, relacionan erróneamente lo justo y lo bueno. No intentaremos dar forma a nuestra vida atendiendo primero al bien, independientemente definido. No es nuestro propósito el de revelar principalmente nuestra naturaleza, sino más bien los principios que admitiríamos que gobernasen las condiciones básicas en que han de formarse estos propósitos y la manera en que deben perseguirse. *Porque el «yo» es anterior a los fines que mediante él se afirman*; incluso un fin dominante tiene que ser elegido entre numerosas posibilidades. [...] Invertiríamos, pues, la relación entre lo justo y lo bueno propuesta por las doctrinas teleológicas y consideraríamos lo justo como prioritario. La teoría moral se desarrolla, entonces, actuando en sentido contrario. [Los subrayados son míos.] (619)

La prioridad del «yo» sobre sus fines significa que no soy meramente el receptáculo pasivo de los intereses, atributos y propósitos acumulados arrojados por la experiencia; no soy simplemente el producto de los caprichos de las circunstancias, sino siempre, irreductiblemente, un agente activo, discrecional, que puedo diferenciarme de mi entorno, y capaz de elección. Identificar cualquier conjunto de características como *mis* intereses, ambiciones, deseos, etc., implica siempre algún sujeto «yo» ubicado detrás de ellos, y la forma de este «yo» debe estar dada con anterioridad a cualquiera de los fines o atributos que detento. Como expresa Rawls, «incluso un fin dominante tiene que ser elegido entre numerosas posi-

bilidades». Y antes de que se pueda seleccionar un fin, debe existir un «yo» que lo haga.

Pero, ¿en qué sentido exactamente el «yo», en tanto un agente de elección, «debe» ser lógicamente anterior a los fines que elige? Un sentido de la prioridad es el «deber» moral que refleja el imperativo de respetar sobre todo la autonomía del individuo, de considerar al ser humano como el portador de una dignidad que está más allá de los roles que éste asuma y de los fines que pueda perseguir. Pero hay otro sentido en el cual el yo «debe» ser lógicamente anterior a los fines que sostiene —anterior en el sentido de independientemente identificable— y ésta es una exigencia epistemológica.

En este punto la consideración del «yo» reproduce las perplejidades que encontramos en el caso de la justicia. En el primero, necesitábamos un punto de vista evaluativo independiente de los valores sociales que prevalezcan. En el caso de la persona, necesitamos una noción de sujeto independiente de sus necesidades y objetivos. Como la prioridad de la justicia surgía a partir de la necesidad de distinguir el criterio de evaluación de la sociedad que estaba siendo evaluada, la prioridad del «yo» surge a partir de la necesidad paralela de distinguir al sujeto de su situación. Por más que Rawls mismo no ofrezca esta consideración, creo que está implícita en su teoría, y es una reconstrucción razonable de las perplejidades que pretende analizar.

Si el «yo» no consistiese en nada más que en una concatenación de diversos deseos, necesidades y fines contingentes, no habría ningún modo no arbitrario —ya sea para ese «yo» o para un observador externo— de identificar *determinados* deseos, intereses y fines como característicos de cualquier sujeto en particular. Antes que pertenecer al sujeto, serían el sujeto. Pero no sería posible discernir entre el sujeto que serían y el torrente de atributos no diferenciados de una situación inarticulada, lo que implica afirmar que no existiría en absoluto un sujeto, al menos ningún sujeto que podamos reconocer o identificar como parecido a una persona.

Cualquier teoría del «yo» expresada en la fórmula «Yo soy *x*, *y*, y *z*», en lugar de la forma «Yo tengo *x*, *y*, y *z*», (siendo *x*, *y*, y *z* deseos, etc.) desbarata la distancia entre sujeto y situación, distancia que es necesaria para cualquier concepción coherente de un sujeto humano en particular. Este espacio, o grado de distanciamiento, resulta esencial al aspecto ineludiblemente

posesivo de cualquier concepción coherente del «yo». El aspecto posesivo del «yo» significa que nunca puedo resultar completamente conformado por mis atributos; que siempre existen algunos atributos que debo tener y no ser. De lo contrario, cualquier cambio en mi situación, no importa cuán sutil pueda ser, cambiaría la persona que soy. Pero considerado literalmente, y dado que mi situación cambia, al menos en algún aspecto, en cada momento que transcurre, esto significaría que mi identidad se confundiría indistinguiblemente con «mi» situación. Si no se traza alguna distinción entre el sujeto y el objeto de la posesión, se torna imposible distinguir lo que soy de lo que es mío, y esto nos deja con lo que podríamos denominar un sujeto *radicalmente situado*.

Ahora bien, un sujeto radicalmente situado no se adecua a la noción de persona, de la misma manera que un criterio de evaluación completamente implicado por los valores vigentes no se adecua a la noción de justicia. El impulso hacia la prioridad que se refleja en la búsqueda de un punto arquimédico es la respuesta a ambos predicamentos.

Pero en ambos casos, las alternativas están seriamente acotadas, las posibles ubicaciones del punto arquimédico se ven severamente restringidas. En el caso de la justicia, la alternativa a una concepción situada parecería una apelación más allá de la experiencia, o sea, a principios a priori. Pero esto sería afirmar la prioridad deseada de manera extrema, y sólo lograría el distanciamiento necesario al precio de la arbitrariedad. Una dificultad similar surge en el caso del sujeto. Ya que un sujeto separado totalmente de sus aspectos empíricamente dados no sería más un tipo de conciencia abstracta (¿consciente de qué?): un sujeto radicalmente situado que se ve reemplazado por un sujeto radicalmente no corporeizado. Nuevamente aquí «necesitamos una concepción que nos permita contemplar nuestros objetivos desde lejos», pero no desde tan lejos que nuestra meta desaparezca de nuestra perspectiva y que nuestra visión se disuelva en la abstracción.

De modo que, al menos en líneas generales, podemos ver cómo el argumento aglutina y relaciona la primacía de la justicia, el rechazo de la teleología y la prioridad del «yo», y, finalmente, cómo estos postulados prestan apoyo a posiciones liberales tradicionales. La conexión entre la perspectiva metaética y la concepción del «yo» puede percibirse en el hecho de que las concepciones deontológicas y teleológicas dan cuenta

de la unidad del «yo» de distintas maneras. De acuerdo con Rawls, para quien las concepciones teleológicas suponen que la unidad del «yo» se alcanza con el curso de la experiencia —en el caso del hedonismo, maximizando la cantidad de experiencias placenteras dentro de sus «límites psíquicos», (620)—, la justicia como imparcialidad revierte esta perspectiva y concibe la unidad del «yo» como algo previamente establecido, formado con anterioridad a las elecciones que se hacen en el curso de la experiencia.

> Los individuos [en la posición original] consideran la personalidad moral, y no la capacidad de placer y de dolor, como el aspecto fundamental del «yo». [...] La idea principal consiste en que, dada la prioridad de lo justo, la elección de nuestra concepción de lo bueno se estructura dentro de unos límites definidos [...] *La unidad esencial del «yo» es facilitada ya por la concepción de lo justo.* [Los subrayados son míos.] (621-622)

Aquí, como en el caso de la igualdad de libertades y la prioridad de lo justo a las que nos referimos anteriormente, se puede ver que los postulados deontológicos generan conclusiones liberales conocidas y las fundamentan con mayor solidez de lo que permitirían el empirismo tradicional o la metafísica utilitarista. El tema común a una gran parte de la doctrina liberal clásica que surge a partir de la consideración deontológica de la unidad del «yo» es la noción del sujeto humano como un agente de elección soberano, una criatura cuyos fines son elegidos antes que dados, que llega a sus propósitos y objetivos por actos de voluntad, en contraste, digamos, con los actos de cognición. «Así, una persona moral es un sujeto con fines que él ha elegido, y su preferencia fundamental se inclina a favor de las condiciones que le permitan construir un modo de vida que exprese su naturaleza de ente racional, libre e igual, tan plenamente como las circunstancias lo consientan» (620).

La unidad antecedente del «yo» significa que el sujeto, no importa cuán condicionado por su entorno se encuentre, siempre es irreductiblemente anterior a sus valores y fines, y nunca completamente construido por ellos. Aunque puedan darse circunstancias en las cuales las condiciones sean formidables y las elecciones pocas, la acción soberana del hombre como tal no depende de ninguna condición particular de existencia, sino que está garantizada por adelantado. En la concepción

deontológica, nunca podríamos estar directamente condicionados de tal manera que nuestro yo llegase a ser completamente construido por nuestra situación, ni nuestros fines completamente determinados de manera tal que el «yo» dejara de ser anterior a ellos. Las consecuencias para la política y la justicia de una perspectiva de este tipo son considerables. Si se asume que el hombre es por naturaleza un ser que elige sus fines antes que un ser que los descubre (como lo concebían los antiguos), su preferencia fundamental debe inclinarse necesariamente por las condiciones de elección antes que, digamos, a favor de las condiciones de autoconocimiento.

La fuerza cierta del postulado de la primacía de la justicia, en su dimensión moral y epistemológica, puede apreciarse más claramente a la luz del debate sobre el «yo» que propone Rawls. Puesto que el «yo» debe su constitución, su condición de antecedente, al concepto de lo justo, sólo podemos expresar nuestra verdadera naturaleza cuando actuamos a partir de un sentido de la justicia. Es por esto que el sentido de la justicia no puede considerarse meramente como un deseo entre otros, sino que debe estimarse como una motivación perteneciente a un orden cualitativamente superior: la justicia no es meramente un valor importante entre otros sino verdaderamente la primera virtud de las instituciones sociales.

> [...] El deseo de expresar nuestra naturaleza como un ser racional, libre e igual, sólo puede realizarse actuando sobre la base de que los principios de lo justo y de la justicia tengan la primacía. [...] Es actuando desde esta superioridad como se expresa nuestra libertad de la contingencia y de la casualidad. Por lo tanto, para realizar nuestra naturaleza, no tenemos más alternativa que la de proteger nuestro sentido de la justicia para que rija nuestros restantes objetivos. Este sentimiento no puede realizarse, si está comprometido y equilibrado frente a otros fines, como un solo deseo entre los demás. Es un deseo de conducirse de un cierto modo, sobre todo lo demás, un esfuerzo que contiene en sí mismo su propia prioridad. (634)
>
> Lo que no podemos hacer es expresar nuestra naturaleza siguiendo un proyecto que considera el sentido de la justicia sólo como un deseo que ha de valorarse frente a otros. Porque este sentimiento revel lo que la persona es, y transigir en él no es alcanzar para el «yo» el reino de la libertad, sino dar paso a las contingencias y a los accidentes del mundo. (635)

La conexión entre la primacía de la justicia y otros rasgos centrales de la concepción de Rawls —por ejemplo la prioridad de lo justo, la metaética deontológica y la unidad antecedente del «yo»— revela algo acerca de la estructura de la teoría global de Rawls, e indica qué tan profundo y poderoso se propone ser el postulado de la primacía de la justicia. También sugiere la manera en que este conjunto de afirmaciones, si puede defenderse, proporciona un fundamento sólido, moral y epistemológico a la vez, para ciertas doctrinas liberales centrales. Nos hemos propuesto comprender estos postulados y aclarar sus vinculaciones al considerarlos como respuestas a las perplejidades planteadas por dos reconstrucciones relacionadas. La primera apunta a un criterio de evaluación que no esté comprometido ni con criterios vigentes ni con criterios dados arbitrariamente; la segunda pretende una consideración del «yo» que no sea del tipo «radicalmente situado», y en consecuencia inextricable de su entorno, ni del tipo «radicalmente no corporeizado», y por lo tanto puramente formal. Cada una de las reconstrucciones establece un conjunto de alternativas inaceptables y requiere para su solución una clase de punto arquimédico que sea capaz de separarse de lo contingente sin, al mismo tiempo, incurrir en la arbitrariedad.

En relación con este punto, el proyecto de Rawls se asemeja bastante al proyecto kantiano. Pero a pesar de sus afinidades deontológicas y su agenda aproximadamente común, la solución propuesta por Rawls se aparta radicalmente de la solución de Kant. La diferencia refleja la preocupación de Rawls por establecer las prioridades deontológicas requeridas, incluyendo la prioridad del «yo», sin recurrir a un sujeto trascendente ni en ningún sentido incorpóreo. Este contraste asume un interés especial puesto que el idealismo kantiano —la dimensión que Rawls intenta evitar por todos los medios— llevaría a buena parte de la filosofía continental de los siglos XIX y XX hacia una dirección en gran medida extraña a la tradición angloamericana del pensamiento político y moral, en la cual el trabajo de Rawls está firmemente instalado.

Para Kant, la prioridad de lo justo (o la supremacía de la ley moral), y la unidad del «yo»(o la unidad sintética de la percepción), sólo podrían establecerse por medio de una deducción trascendental y de la formulación de un dominio inteligible o noumenal como el presupuesto necesario de nuestra capaci-

dad de actuar libremente y de acceder al autoconocimiento. Rawls rechaza la metafísica kantiana, pero cree que puede preservar su fuerza moral «dentro del alcance de una teoría empírica» (Rawls 1979: 18). Éste es el papel que desempeña la posición original.

Liberalismo sin metafísica: la posición original

La posición original es la respuesta que Rawls formula a Kant; es su alternativa al camino representado por la *Crítica de la Razón Pura* y la clave para la solución que propone a las perplejidades que hemos analizado. Es la posición original lo que «nos permite contemplar nuestros objetivos desde lejos» pero no tan lejos como para dejarnos en el dominio de lo trascendente. Rawls intenta satisfacer estas demandas con la descripción de una situación inicial de imparcialidad, y con la definición de aquellos principios que acordarían como justas las partes racionales sujetas a las condiciones de dicha situación.

La posición original cuenta con dos ingredientes cruciales para resolver los dilemas descriptos por la reconstrucción, y para responder a la necesidad de un punto arquimédico. Cada uno toma la forma de un supuesto sobre las partes de la posición original: el primero establece lo que éstas no conocen, el segundo lo que en efecto conocen. No disponen de información alguna que les permita distinguir entre ellas y cualquier otra, en tanto seres humanos particulares que son. Éste es el supuesto del velo de la ignorancia. Significa que se asume que las partes no tienen ningún conocimiento relativo a su lugar en la sociedad; su raza, sexo o clase su salud o fortuna; su inteligencia su fuerza u otras habilidades y ventajas naturales. Ni siquiera conocen sus concepciones de lo bueno, sus valores, objetivos, o propósitos para la vida. Saben que de hecho sí poseen tales concepciones y opinan que vale la pena intentar promoverlas, sean las que sean, pero deben elegir los principios de justicia en una situación de ignorancia temporaria de tales concepciones. El objetivo de esta restricción es impedir que la elección de principios se vea influida por la contingencia de circunstancias naturales y sociales, y abstraerla de todas las consideraciones que se juz-

gan como irrelevantes desde un punto de vista moral. Es el velo de la ignorancia lo que asegura que los principios de justicia serán elegidos bajo condiciones de igualdad y equidad. Puesto que las partes en el contrato no se distinguen por intereses diferentes, una consecuencia agregada del velo de la ignorancia es garantizar que el acuerdo inicial sea unánime.

Lo que las partes sí saben es que ellas, como todas las demás, valoran ciertos bienes sociales primarios. Bienes primarios son «cosas que se supone que un hombre racional desea, sin importar si desea cualquier otra», e incluyen cosas tales como derechos y libertades, oportunidades y potestades, ingresos y riquezas. Más allá de los valores, planes y objetivos últimos de una persona, se considera que existen ciertas cosas de las cuales preferiría tener más en lugar de menos, a partir de la base de que probablemente serán útiles para promover todos los fines, sean cuales sean tales fines. De modo que mientras las partes en la posición original son ignorantes de sus fines particulares, se asume que todas ellas son motivadas por el deseo de ciertos bienes primarios.

El contenido preciso de la lista de los bienes primarios está dado por lo que Rawls denomina la teoría débil del bien. Es débil en el sentido de que incorpora supuestos mínimos y ampliamente compartidos sobre la clase de cosas que probablemente resulten útiles para todas las concepciones particulares de lo bueno, y por lo tanto, sean probablemente compartidas por las personas más allá de las diferencias entre sus deseos más específicos. La teoría débil del bien se distingue de la teoría completa del bien en que la primera no puede proporcionar ninguna base para juzgar o elegir entre distintos valores o fines particulares. Por lo tanto, aunque el velo de la ignorancia asegura que las partes deliberen en condiciones de equidad y unanimidad, la consideración de los bienes primarios genera la motivación mínima necesaria para poner en movimiento un problema de elección racional, y para hacer posible una determinada solución. Los dos supuestos combinados aseguran que las partes actúen solamente sobre aquellos intereses que resultan comunes, es decir, los compartidos por todas las personas racionales, el principal de los cuales consiste en el interés de establecer términos de cooperación social de manera tal que cada persona tenga la libertad más absoluta para llevar a cabo sus objetivos y propósitos en tanto sean compatibles con una libertad igual para los demás.

Los principios de la justicia surgen a partir de la posición original, a través de lo que puede considerarse como un procedimiento en tres etapas. Primero surge la teoría débil del bien, representada por la descripción de la situación de la elección inicial. A partir de la teoría débil se derivan los dos principios de justicia, que a su vez definen el concepto del bien y proporcionan una interpretación de valores tales como el bien de la comunidad. Es importante tener en cuenta que aun cuando la teoría débil del bien es anterior a la teoría de lo justo y los principios de la justicia, no es una teoría lo suficientemente substancial como para debilitar la prioridad de lo justo sobre lo bueno que otorga su carácter deontológico a la concepción. La prioridad de lo justo en la que se basa la teoría se establece con respecto a la teoría completa del bien, la cual se relaciona con valores y fines particulares. La teoría general del bien sólo aparece después de los principios de la justicia y a la luz de ellos. Como explica Rawls:

> Pero, para establecer estos principios [de la justicia], es necesario contar con alguna noción de bondad, porque necesitamos hipótesis acerca de los motivos de los individuos en la situación original. Como estas hipótesis no deben poner en peligro la posesión prioritaria del concepto de lo justo, la teoría del bien utilizada para argüir a favor de los principios de la justicia se reduce a lo simplemente indispensable. Yo llamo a esta descripción del bien la teoría débil: su propósito es el de asegurar las premisas acerca de los bienes primarios, requeridas para alcanzar los principios de la justicia. Una vez elaborada esta teoría y explicados los bienes primarios, nos hallaremos en condiciones de emplear los principios de la justicia en el ulterior desarrollo de lo que llamaré la teoría completa del bien. (438)

Este procedimiento en tres etapas parece cumplir con los requisitos deontológicos de Rawls de la siguiente manera: la prioridad de lo justo sobre (la teoría completa de) el bien satisface la exigencia de que el criterio de evaluación sea anterior a los objetos que evalúa y discernible de ellos, no comprometido por implicación en los deseos y objetivos existentes. Y el hecho de que los principios de lo justo no provengan de ningún lugar sino de una teoría débil del bien relacionada con deseos humanos reales (aunque muy generales), otorga a los principios una base determinada y evita que sean arbitrarios y aislados del mundo. Y de esta manera, sin

recurrir a las deducciones trascendentales, parece posible encontrar un punto arquimédico que no está ni radicalmente situado ni es radicalmente incorpóreo; ni «a merced de los deseos e intereses presentes» ni dependiente de consideraciones apriorísticas.

> En resumen, el punto esencial es el de que, a pesar de los rasgos individualistas de la justicia como imparcialidad, *los dos principios de la justicia no son contingencias que actúan sobre los deseos o las condiciones sociales presentes*. Así, somos capaces de deducir una concepción de la estructura básica justa, y un ideal de la persona compatible con ella, que puede servir como ejemplo para las instituciones, y para enfocar la dirección del cambio social. Para encontrar un punto de Arquímedes *no es necesario recurrir a principios perfeccionistas o apriorísticos*. Dando por supuesto ciertos deseos generales, tales como el deseo de bienes sociales primarios, y tomando como base los acuerdos que se obtendrían en una situación inicial claramente definida, podemos lograr la independencia necesaria de las actuales circunstancias. [Los subrayados son míos.] (300)

Este es, en resumen, el procedimiento por el cual se derivan los dos principios de justicia. Como destaca Rawls, la justicia como imparcialidad, al igual que otras perspectivas contractuales, está conformada por dos partes. La primera es una interpretación de la situación inicial y del problema de la elección que se establece en ésta; la segunda involucra los dos principios de justicia en los que se postula que existiría acuerdo. «Se puede entonces aceptar la primera parte de la teoría (o una variante de la misma) sin aceptar la otra, y viceversa» (33). Aun si no determinamos los principios de justicia que de hecho Rawls cree que se seleccionarían, es posible identificar dos tipos de objeciones que probablemente surjan en la transición desde la primera mitad de la teoría hacia la segunda.

Una primera visión crítica cuestionaría si la posición original logra un genuino distanciamiento con respecto a los deseos y objetivos existentes. Este tipo de objeción probablemente se concentraría en la consideración de los bienes primarios o en algún otro aspecto de la teoría débil del bien, y afirmaría que está sesgada en favor de ciertas concepciones particulares de lo bueno y en contra de otras. Podría impugnar el postulado que propone Rawls, en el sentido de que la lista de bienes

primarios tenga realmente un valor equivalente, o casi equivalente, para todos los estilos de vida. Podría cuestionar el carácter de «débil» de esta teoría del bien, afirmando que debilita la imparcialidad de la situación inicial; que introduce presupuestos no compartidos universalmente; que se implica demasiado profundamente en las preferencias contingentes de, digamos, los planes de vida de la burguesía liberal occidental; y que, después de todo, los principios resultantes son el producto de valores prevalecientes.

Otro conjunto de objeciones argumentaría que, por otra parte, la posición original logra demasiada separación respecto de las circunstancias humanas; que la situación inicial que describe es demasiado abstracta como para producir los principios que Rawls afirma que produciría, y en todo caso, cualquier principio determinado. Una objeción de este tipo podría probablemente discrepar con el velo de la ignorancia sobre la base de que excluye información moralmente relevante, necesaria para generar cualquier resultado significativo. Argumentaría que la noción de la persona inmersa en la posición original es demasiado formal y abstracta, demasiado separada de la contingencia como para dar cuenta de las motivaciones necesarias. Mientras que la primera objeción formula la crítica de que la teoría débil del bien es demasiado robusta como para ser justa, la segunda argumenta que el velo de la ignorancia resulta demasiado opaco como para producir una solución determinada.

No me abocaré a ninguna de estas objeciones críticas en esta obra. Dado que nos interesa el proyecto deontológico en su totalidad, nuestro interés en la posición original es más general. En términos simples: si consideramos que la posición original de Rawls es su respuesta a Kant, ¿resulta satisfactoria esta respuesta? ¿Tiene éxito en su aspiración de reformular la moral kantiana y los postulados políticos «dentro del alcance de una teoría empírica»? ¿Puede proporcionar un fundamento al liberalismo deontológico al mismo tiempo que evita el «entorno» metafísicamente cuestionable de la teoría kantiana? Más específicamente, ¿puede la descripción de la posición original acomodar y apoyar el postulado de la primacía de la justicia en el sentido fuerte que Rawls intenta proponer?

Dentro de una lectura de la posición original, una interpretación fuertemente empirista que el mismo Rawls invita a

adoptar, no puede sostenerse el postulado deontológico. Para descubrir a qué se debe esto, debemos examinar sus condiciones más rigurosamente, en parte con la perspectiva de comprender el tipo de postulado que encarnan estas condiciones. Por consiguiente no estaremos interesados en la cuestión de si la teoría débil del bien es demasiado «robusta» o demasiado «débil» como para generar los principios de justicia que Rawls afirma que genera, sino más bien en la cuestión de qué es lo que hace que la teoría resulte demasiado «robusta» o «débil», y en qué manera se relaciona esta consideración con la justificación de qué es lo que *hace* que la justicia sea primaria. Pero quizás lo más indicado sea abocarnos al examen de las condiciones que caracterizan a la posición original tal como las describe Rawls. Para esto debemos enfocar nuestra atención hacia las circunstancias de la justicia.

La justicia y sus circunstancias: las objeciones empiristas

Las circunstancias de la justicia son las condiciones que involucran la virtud de la justicia. Son las condiciones que prevalecen en las sociedades humanas y tornan posible y necesaria la cooperación humana. La sociedad se considera como una empresa cooperativa para lograr beneficios mutuos, lo que significa que está signada de manera característica por el conflicto, tanto como por la identidad de intereses; una identidad de intereses que se expresa en el hecho de que todos quieren ganar a partir de la cooperación mutua, mientras que el conflicto se muestra en el hecho de que, dados sus intereses y fines divergentes, la gente difiere respecto de la manera en que deben distribuirse los beneficios de su cooperación. Los principios son necesarios para especificar los acuerdos a través de los cuales puedan resolverse tales demandas, y el papel de la justicia consiste en proporcionarlos. Las condiciones básicas que hacen necesarios estos acuerdos resolutivos constituyen las circunstancias de la justicia.

De acuerdo con Hume, Rawls señala que estas circunstancias son de dos clases, objetiva y subjetiva. Las circunstancias objetivas de la justicia incorporan hechos tales como la mode-

rada escasez de recursos, mientras que las circunstancias subjetivas tienen que ver con los sujetos de cooperación, particularmente con el hecho de que los intereses y fines que los caracterizan son diferentes. Esto significa que cada persona tiene un plan de vida o concepción del bien diferente que considera vale la pena llevar a cabo. Rawls enfatiza este aspecto asumiendo que, al menos como lo concibe en la posición original, las partes son mutuamente desinteresadas, que les importa promover su propia concepción del bien y no la de alguna otra, y que al fomentar sus fines no están ligadas a los fines de las otras partes por lazos morales anteriores. De esta manera se resumen las circunstancias de la justicia:

> Así, puede decirse brevemente que las circunstancias de la justicia se dan siempre que, en condiciones de escasez moderada, las personas presenten demandas conflictivas ante la división de las ventajas sociales. Si estas circunstancias no existieran, no habría ninguna ocasión para que se diera la virtud de la justicia, del mismo modo que dada la ausencia de amenazas vitales y corporales, no habría oportunidades para que se suscitara el valor físico. (153-154)

Las circunstancias de la justicia son las circunstancias que hacen surgir la virtud de la justicia. En su ausencia, la virtud de la justicia sería vana; ni siquiera sería requerida para esa cuestión aun cuando fuera posible. «Pero una sociedad humana *está* caracterizada por las circunstancias de la justicia» [Los subrayados son míos] (155). En consecuencia la virtud de la justicia *es* necesaria.

Las condiciones que hacen surgir la virtud de la justicia son condiciones empíricas. Sobre esto Rawls es claro y explícito. «La teoría moral debe tener la libertad de utilizar, a su gusto, supuestos contingentes y hechos generales.» No puede proceder de otra manera. Lo que es esencial es que las premisas sean «verdaderas y adecuadamente generales» (71, 187).

> [...] Los principios fundamentales de la justicia dependen en efecto de hechos naturales acerca del hombre en sociedad. Esta dependencia se hace explícita mediante la descripción de la posición original: la decisión de las partes se toma a la luz del conocimiento general. Más aún, los diversos elementos de la posición original presuponen muchas cosas acerca de la circunstancia de la vida humana [...] Si estos supuestos son verdaderos y adecua-

damente generales, todo está en orden; sin estos elementos todo el esquema se hallará vacío y carente de sentido. (187)

Pero desde una comprensión empirista de la posición original parece surgir una profunda contradicción con las demandas deontológicas, ya que si la justicia para su virtud depende de ciertas condiciones previas empíricas, no está claro de qué manera podría afirmarse incondicionalmente su prioridad. Rawls explica que toma prestada su consideración de las circunstancias de la justicia de Hume (152). Pero las circunstancias de Hume no pueden fundamentar la prioridad de lo justo en el sentido deontológico. Después de todo, son condiciones empíricas. Para establecer la primacía de la justicia en el sentido categórico que requiere el postulado de Rawls, debería demostrar no sólo de qué manera prevalecen las circunstancias de la justicia en todas las sociedades, sino también que lo hacen hasta un punto tal que la virtud de la justicia siempre está más completa o ampliamente involucrada que cualquier otra virtud. De lo contrario, sólo estaría autorizado a concluir que la justicia es la primera virtud de cierto tipo de sociedades, específicamente aquéllas cuyas condiciones son tales que la resolución de las demandas en conflicto entre partes mutuamente desinteresadas constituye la prioridad social más urgente.

Indudablemente, un sociólogo podría argumentar, por ejemplo, que dada la creciente escasez de energía y de otros recursos básicos que enfrentan las sociedades industriales desarrolladas, combinada con la quiebra del consenso y la pérdida de propósitos comunes (las circunstancias objetivas y subjetivas, respectivamente), las circunstancias de la justicia han llegado a prevalecer con una intensidad tal que la justicia se ha transformado, para esas sociedades, en la primera virtud. Pero si Rawls postula que la primacía de la justicia depende de una generalización como ésta, necesitaría al menos proporcionar el apoyo sociológicamente relevante. Afirmar solamente que «*una* sociedad humana *está* caracterizada por las circunstancias de la justicia» [Los subrayados son míos] (155), no sería suficiente.

La noción de que la primacía de la justicia podría fundamentarse empíricamente deja de ser del todo plausible cuando consideramos cuán improbable debe ser la generalización necesaria, al menos cuando se aplica a lo largo de todo el aba-

nico de instituciones sociales. Ya que mientras podemos imaginar muy fácilmente que ciertas asociaciones a gran escala como el moderno estado nación podrían satisfacer en muchos casos sus requerimientos, también podemos imaginar sin demasiada dificultad un rango de asociaciones más íntimas o solidarias en las que los valores y objetivos de los participantes coincidan más estrechamente, de forma tal que las circunstancias de la justicia prevalezcan en un grado relativamente reducido. Como Hume mismo observa, no necesitamos recurrir a visiones utópicas o a la ficción de los poetas para imaginar tales condiciones, sino que «podemos descubrir esta verdad mediante la experiencia y observación comunes» (Hume 1739: 495; 721).

> En la disposición actual del corazón humano resultaría, tal vez, difícil encontrar instancias completas de tales afectos ampliados; pero sin embargo podemos observar que el caso de las familias se aproxima; y cuanto más fuerte es la benevolencia mutua entre los individuos, es más lo que se aproxima; hasta que toda distinción de la propiedad se vea en gran medida perdida y confundida entre ellos. Entre las personas casadas, la ley supone que el cemento de la amistad es tan fuerte como para abolir todas las divisiones de sus propiedades; y a menudo tiene en efecto la fuerza que se le atribuye. (Hume 1777: 17-18)

Mientras que la institución de la familia puede representar un caso extremo en este aspecto, podemos fácilmente imaginar una serie de casos intermedios de instituciones sociales, un abanico de asociaciones humanas caracterizadas, en diversos grados, por las circunstancias de la justicia. Estas incluirían, en varios puntos a lo largo del espectro, tribus, vecindades, ciudades, universidades, uniones comerciales, movimientos de liberación nacional y nacionalismos establecidos, y una amplia variedad de comunidades étnicas, religiosas, culturales y lingüísticas con identidades más o menos claramente definidas y propósitos compartidos, precisamente los atributos cuya presencia implica la relativa ausencia de las circunstancias de la justicia. Aun cuando las circunstancias de la justicia bien podrían existir en todos estos casos, probablemente no predominarían, al menos no en un grado tal que la justicia estuviera involucrada en todos los casos en una medida superior a la que pudiera estarlo cualquier otra virtud. Dentro de la interpretación empirista de la posición original, la justicia puede ser primaria sólo para aquellas sociedades acosadas por la su-

ficiente discordia como para hacer que la resolución de los intereses y objetivos en conflicto sea la consideración moral y política dominante; la justicia es la primera virtud de las instituciones sociales no de un modo absoluto, como lo es la verdad para las teorías, sino sólo de manera condicional, como el coraje físico lo es para una zona en guerra.

Pero esta interpretación sugiere un sentido ulterior en el que la primacía de la justicia se ve debilitada por la consideración empirista de las circunstancias de la justicia. Se concentra en el sentido en el que la justicia aparece como una virtud compensatoria, cuya ventaja moral consiste en su capacidad de remediar las condiciones que se han degradado. Pero si la virtud de la justicia se mide por las condiciones moralmente disminuidas que son su condición previa, entonces la *ausencia* de tales condiciones, sin importar cómo pueda describirse este estado de la cuestión, debe encarnar una virtud rival de una prioridad por lo menos conmensurable, la que se ve involucrada en la medida en que la justicia no lo hace. Si el coraje físico es una virtud sólo frente a condiciones perjudiciales, entonces la paz y la tranquilidad que negarían al coraje su ocasión de hacerse presente, deben, seguramente, ser virtudes de una condición al menos equivalente. Y lo mismo sucede con la justicia. Como confirma la consideración de Hume, el carácter compensatorio de la justicia implica otro conjunto de virtudes de orden por lo menos comparable.

> Ya he señalado que la justicia debe su origen a las convenciones humanas, y que éstas se proponen como remedio de algunos inconvenientes debidos a la coincidencia de ciertas *cualidades* de la mente humana con la *situación* de objetos externos. Las cualidades de la mente son el *egoísmo* y la *generosidad limitada*, y la situación de los objetos consiste en su *facilidad de cambio*, unida a su *escasez* en comparación con las necesidades y deseos de los hombres [...] Elévense en grado suficiente la benevolencia de los hombres o el botín de la naturaleza y haréis que la justicia se convierta en algo inútil, supliendo su lugar mediante virtudes mucho más nobles y bienes más valiosos. (Hume 1739: 494-495; 720-721)

Invocar las circunstancias de la justicia significa admitir al mismo tiempo, por lo menos de manera implícita, las circunstancias de la benevolencia, de la fraternidad o de sentimientos parecidos, cualquiera que sea su descripción posible; tales son

las circunstancias que prevalecen en la medida en que la justicia no lo hace, y la virtud a la que otorgan definición debe ser una virtud de una condición al menos correlativa.

Una consecuencia del aspecto compensatorio de la justicia es que no podemos decir anticipadamente si, en cualquier instancia particular, un incremento en la justicia se asocia con una mejora moral total. Esto se da debido a que un aumento de la justicia puede producirse de alguna de dos maneras: puede surgir donde antes había injusticia, o puede aparecer donde antes no había ni justicia ni injusticia sino una medida suficiente de benevolencia o fraternidad tal que la virtud de la justicia no había sido involucrada ampliamente. Donde la justicia reemplaza a la injusticia, permaneciendo lo demás constante, la mejora moral total resulta evidente. Por otra parte, en el caso de un incremento de la justicia que refleje alguna transformación en la calidad de motivaciones y disposiciones preexistentes, podría bien tratarse de una disminución del balance moral total.

Cuando la fraternidad se desvanece, se puede hacer más justicia, pero puede necesitarse aún más para volver al mismo *status quo* moral. Además, no hay garantías de que la justicia y sus virtudes rivales sean perfectamente conmensurables. La ruptura de ciertos vínculos personales y cívicos puede representar una pérdida moral que ni siquiera una considerable cantidad de justicia pudiera reparar. ¿Podemos dar por sentado que una rasgadura en el tejido de entendimientos y compromisos implícitos logra una restauración moral completa en tanto todos «hacen lo que deben» de ahí en adelante?

Consideremos, por ejemplo, una situación conocida más o menos ideal, donde las relaciones se rigen en gran medida por el afecto espontáneo, y donde, por lo tanto, las circunstancias de la justicia prevalecen en un grado relativamente menor. Los derechos individuales y los procedimientos de justicia imparcial son raramente invocados, no porque la injusticia sea predominante sino porque su interés está superado por un espíritu de generosidad en el que pocas veces me siento inclinado a exigir mi justa parte. Tampoco esta generosidad implica necesariamente que reciba por benevolencia una parte que sea igual o mayor a la que estaría autorizado a recibir bajo principios imparciales de justicia. Puedo obtener menos. El punto no es que obtengo lo mismo que obtendría de otra manera, sólo que más espontáneamente, sino simplemente que

las cuestiones respecto de qué es lo que obtengo y de a qué estoy obligado no cobran demasiada importancia en el contexto total de este modo de vida.

Ahora bien, imaginemos que un día la armoniosa familia se ve atrapada en la discordia. Los intereses crecen de manera divergente y las circunstancias de la justicia se hacen más agudas. El afecto y espontaneidad de los días anteriores dan lugar a demandas de imparcialidad y de la observancia de derechos. Y posteriormente imaginemos que la antigua generosidad es reemplazada por un ánimo sensato de integridad sin excepciones, y que las nuevas necesidades morales se cumplen con una cantidad máxima de justicia, de forma tal que ninguna injusticia prevalezca. Los padres e hijos equilibrados reflexivamente, aceptan respetuosamente, aunque con renuencia, guiarse por los dos principios de la justicia, e incluso logran realizar las condiciones de estabilidad y congruencia de modo tal que se logra el bien de la justicia dentro de su familia. Ahora bien, ¿qué tenemos que deducir de esto? ¿Estamos preparados para decir que la llegada de la justicia, no importa cuál sea su dimensión, restituye su carácter completamente moral a la situación, y que la única diferencia, es una diferencia psicológica? Consideremos nuevamente el ejemplo del coraje físico. Imaginemos una sociedad una vez tranquila pero con poco coraje (no por cobardía sino por tranquilidad), y que ahora llegó a ser violenta e inestable, pero donde la virtud del coraje se exhibe en forma destacada, e incluso generalizada. ¿Resulta obvio que preferiríamos la segunda situación desde un punto de vista moral?

Sin duda las inconmensurabilidades, si existen, también podrían guiarnos en la dirección opuesta. Sería posible que, a pesar de la crueldad de las circunstancias del coraje, exista una cierta nobleza que florezca con este nuevo modo de vida, y algo que el espíritu humano no podría alcanzar bajo condiciones más protegidas y que no se vería compensado ni siquiera en la paz más dichosa. Y si la defunción de la *Gemeinschaft* familiar o comunitaria no refleja el comienzo de la mezquindad material sino el florecimiento de la diversidad, o el hecho de que los hijos superan los modos de vida locales del hogar paterno, podríamos estar inclinados a ver el advenimiento de la justicia de un modo más favorable. El punto principal subsiste. Es posible que un aumento de la justicia no esté de hecho asociado con una mejora moral general por lo menos de dos

maneras: ya sea por fracasar completamente en cumplir con el incremento en las circunstancias de la justicia; o por una incapacidad, por más completa que sea, de compensar la pérdida de ciertas «virtudes más nobles y bienes más valiosos».

Si un incremento en la justicia no implica necesariamente una mejora moral incondicional, puede demostrarse también que en algunos casos, la justicia no es una virtud sino un vicio. Esto puede verificarse considerando lo que podríamos denominar la dimensión reflexiva de las circunstancias de la justicia. La dimensión reflexiva se refiere al hecho de que lo que las partes conocen sobre su condición es a su vez un elemento de su condición. Rawls reconoce este aspecto cuando escribe «[...] asumiré, por supuesto, que las personas en la posición original saben que se darán estas circunstancias de la justicia [...]» (154).

Las circunstancias de la justicia, y más específicamente el aspecto subjetivo de estas circunstancias, consisten en parte en las motivaciones de los participantes y en el modo en que perciben sus motivaciones. Si un día las partes llegaran a considerar de manera diferente sus circunstancias, si llegaran a creer que las circunstancias de la justicia (o de la benevolencia) se dieran en una medida mayor o menor que antes, esta verdadera transformación implicaría un cambio en dichas circunstancias. Como Rawls señala en su discusión sobre el bien de la justicia, la conducta conforme a un sentido de justicia puede ser contagiosa; refuerza los supuestos que presupone, e intensifica su propia estabilidad estimulando y afirmando motivaciones similares en otros.

¿Pero cuál es el efecto de este «contagio» cuando se aplica a una situación en donde no se dan las circunstancias de la justicia (o en la medida en que éstas no se dan)? Cuando actúo de acuerdo con un sentido de justicia en circunstancias inapropiadas, digamos en circunstancias en las que se comprometen de manera relevante más bien las virtudes de benevolencia y fraternidad antes que la justicia, mi acto puede no ser meramente superfluo, sino que podría contribuir a una reordenación de las motivaciones y opiniones prevalecientes, transformando con ello las circunstancias de la justicia en cierto grado. Y esto incluso puede ser verdadero en el caso en que el «acto» que realizo de acuerdo con la justicia sea «el mismo acto» que habría realizado de acuerdo con la benevolencia o la fraternidad, sólo que bajo un espíritu diferente. Como en la con-

sideración de la estabilidad que hace Rawls, mi acto y el sentido de justicia que lo informa tienen el efecto de autocumplimiento: originan las condiciones bajo las cuales *hubieran* sido apropiados. Pero en el caso del acto inapropiado de justicia, el resultado es lograr que las circunstancias de la justicia ejerzan más presión sin evocar necesariamente un incremento en la incidencia de la justicia en similar medida.

Los despliegues gratuitos de coraje físico en condiciones de tranquilidad pueden resultar invasores de la misma tranquilidad que no llegan a apreciar y que muy posiblemente sean incapaces de reemplazar. Ocurre lo mismo con la justicia. Si, a resultas de un equivocado sentido de la justicia, un viejo amigo muy cercano insiste repetidamente en calcular y pagar su parte exacta de todos los gastos comunes, o no quiere aceptar ningún favor u hospitalidad salvo bajo la mayor protesta o desconcierto, no sólo me sentiré obligado a ser recíprocamente escrupuloso sino que en el mismo instante puedo comenzar a preguntarme si no habré comprendido de manera equivocada nuestra relación. Las circunstancias de la benevolencia habrán disminuido en este punto, y habrán crecido las circunstancias de la justicia. Esto se deriva de una consecuencia de la dimensión reflexiva de (los aspectos subjetivos de) las circunstancias de la justicia. Pero como ya hemos visto, no hay garantías de que el nuevo sentido de justicia pueda reemplazar completamente a la anterior espontaneidad, aun en aquellos casos donde no se produce ninguna injusticia. Puesto que el ejercicio de la justicia en condiciones inapropiadas habrá originado una decadencia del carácter moral de la asociación en su conjunto, la justicia en este caso no habrá sido una virtud sino un vicio.

De modo que las circunstancias de la justicia no se ajustan bien a la primacía de la justicia y a los temas deontológicos relacionados que Rawls aspira a defender. Dado el contraste entre los dos linajes filosóficos de ambos enfoques, no es de extrañar que surjan inconsistencias. Donde las circunstancias de la justicia siguen explícitamente a Hume, «la explicación que de ellas hace Hume es particularmente lúcida y el resumen anterior no añade nada esencial a los análisis más completos hechos por él» (153); la concepción deontológica en el núcleo de la teoría de Rawls encuentra su formulación primaria en Kant, cuya epistemología y ética en gran parte estaban

dirigidas contra la tradición empirista y utilitarista que Hume representa. Para Kant, la noción deontológicamente dada de lo justo, que Rawls trata de recuperar, deriva su fuerza de una metafísica moral que impide el recurso a las circunstancias humanas contingentes, en las que Hume basó su consideración de la virtud de la justicia.

Para Hume, la justicia es un producto de las convenciones humanas y «deriva su existencia enteramente de su *uso* necesario para el intercambio y el estado social de la humanidad».

> Así, las reglas de la equidad o la justicia dependen enteramente del estado y condición particulares en los que se ubican los hombres, y deben sus orígenes y existencia a esta utilidad, que es el resultado para el público de su observancia estricta y regular. Reviértase, en cualquier circunstancia considerable, la condición de los hombres; prodúzcase abundancia extrema o extrema necesidad; implántese en el corazón humano la perfecta moderación y humanidad, o una perfecta rapacidad y malicia; al volver la justicia totalmente *inútil*, se destruye totalmente su esencia, y se suspende su obligatoriedad para la humanidad. (Hume 1777: 20)

Por el contrario, para Kant:

> Los *principios empíricos* no sirven nunca para fundamento de leyes morales. Pues la universalidad con que deben valer para todos los seres racionales sin distinción, la necesidad práctica no condicionada que por ello les es atribuida desaparece cuando el fundamento de ella se deriva de la *peculiar constitución de la naturaleza humana* o de las circunstancias contingentes en que se coloca. (Kant 1785: 53)

Si, como parece, una consideración humeana de las circunstancias de la justicia no puede ni sostener ni acomodar la condición privilegiada de la justicia y de lo justo que requiere Rawls y que se deriva a partir de Kant, surge naturalmente la pregunta de por qué el autor no adopta en su lugar una consideración kantiana de la justicia. La respuesta es que Kant, estrictamente hablando, no ofrece ninguna, al menos ninguna que coloque la virtud de la justicia en circunstancias características de la sociedad humana. Tampoco resulta obvio que pueda proporcionar alguna. Hacerlo sería contradecir el punto esencial de la ética kantiana: el hombre actúa moralmente sólo en cuanto es capaz de alzarse por encima de las influen-

cias heterónimas y las determinaciones contingentes de sus condiciones naturales y sociales, y actuar de acuerdo a un principio dado por la razón práctica pura. Para Kant las circunstancias de la justicia no consisten en aquellas condiciones de la sociedad humana que hacen necesaria la justicia, sino más bien en un dominio ideal abstraído a partir de la sociedad humana que hace posible a la justicia y a la moralidad en general. Este territorio es el reino de los fines. Es un reino por encima del mundo fenoménico, como Kant reconoce «ciertamente es sólo un ideal», donde los seres humanos son admitidos no como residentes permanentes sino más probablemente como visitantes fugaces. La admisión a *estas* circunstancias de la justicia no es una condición previa de la virtud moral sino una medida de su cumplimiento, un lugar donde los seres humanos llegan sólo en tanto son capaces de actuar en conformidad con la ley moral dada autónomamente, esto es, sólo en tanto puedan abstraer la voluntad de su situación y actuar como seres no situados, desde un punto de vista universal. Por esto el imperativo categórico puede ordenar sólo que el hombre actúe *como si* fuera un miembro legislador del reino de los fines.

> [R]esultará que, si prescindimos de las diferencias personales de los seres racionales y asimismo de todo contenido de sus fines privados, podrá pensarse un todo de todos los fines (tanto de los seres racionales como fines en sí, como también de los propios fines que cada cual puede proponerse) en enlace sistemático; es decir, un reino de los fines que es posible según los ya citados principios. (Kant 1785: 47)

> Tal reino de los fines sería realmente realizado por máximas, cuya regla prescribe el imperativo categórico a todos los seres racionales, *si éstos universalmente siguieran esas máximas*. Pero aunque el ser racional no pudiera contar con ello, porque él mismo siguiera puntualmente esa máxima, por eso todos los demás habrían de ser fieles a la misma [...] sin embargo, aquella ley: «obra por máximas de un miembro legislador universal en un posible reino de los fines», conserva toda su fuerza, porque manda categóricamente. (Kant 1785: 51)

Como ya hemos visto, Rawls se separa de Kant cuando entran en juego en su concepción los dominios ideales y los sujetos trascendentales, y la consideración del reino de los fines es

uno de tales puntos. Rawls encuentra que dicha noción es insatisfactoria como base para la justicia humana, puesto que parece aplicarse a seres humanos sólo en tanto son separados de sus circunstancias humanas reales, que es lo mismo que decir sólo en tanto dejan de ser seres humanos. Tal noción es, por lo menos, oscura, y puede quedar sujeta a la crítica contra los criterios a priori de evaluación y las concepciones de un «yo» radicalmente incorpóreo. Rawls expresa estas inquietudes, especialmente las concernientes al problema de la arbitrariedad, sugiriendo con Sidgwick que el requisito de abstracción de todas las contingencias por parte de la doctrina de Kant podría no ser capaz de distinguir entre las vidas del santo y del villano, en tanto ambas son vividas según un conjunto consistente de principios elegidos libremente y que se utilizan conscientemente para actuar de acuerdo con ellos. La elección de un «yo» noumenal podría ser —de hecho necesariamente— arbitrario en este sentido. «Kant no afirma que al actuar partiendo de la ley moral se exprese nuestra naturaleza por medios identificables y que actuando sobre principios contrarios, no se exprese» (291). Esta crítica refleja la diferencia más general entre Rawls y Kant respecto del papel de lo empírico y lo apriorístico en la teoría moral, en particular el punto de vista de Rawls de que «el análisis de los conceptos morales y el concepto *a priori*, aunque sean entendidos tradicionalmente, constituyen una base demasiado frágil» para una teoría sustantiva de la justicia. «La teoría moral debe tener la libertad de utilizar, a su gusto, supuestos contingentes y hechos generales» (71).

Para superar estas dificultades, preservando al mismo tiempo la prioridad de lo justo, Rawls pretende reformular la noción del reino de los fines de una manera que permita una consideración empírica de las circunstancias de la justicia, pero que excluya las diferencias contingentes entre personas que de lo contrario prevalecerían.

> La descripción de la posición original interpreta el punto de vista de los seres noumenales, acerca de lo que significa ser un individuo libre y racional. Nuestra naturaleza como tales seres se manifiesta cuando actuamos a partir de principios que elegiríamos cuando esta naturaleza se refleja en las condiciones que determinan la elección. Así los hombres exhiben su libertad, su independencia de las contingencias de la naturaleza y de la sociedad, actuando por medios que reconocerían en la posición original. (292)

La posición original (con el velo de la ignorancia), al igual que el reino de los fines, tiene el efecto de «abstraerse de las diferencias personales entre seres racionales, y además, de todo el contenido de sus fines privados». Pero de manera contraria a Kant, tiene la ventaja adicional de aplicarse a seres humanos reales sujetos a las condiciones normales de la circunstancia humana.

> La posición original puede ser considerada como una interpretación procesal de la concepción kantiana de la autonomía y del imperativo categórico, dentro del sistema de una teoría empírica. Los principios reguladores del reino de los fines son aquéllos que serían elegidos en esta posición, y la descripción de esta situación nos capacita para explicar el sentido según el cual, actuando a partir de estos principios, expresamos nuestra naturaleza de seres libres y racionales. *Estas nociones ya no son puramente trascendentales y carentes de conexión con la conducta humana, ya que la concepción procesal de la posición original nos permite establecer estos lazos.* [Los subrayados son míos.] (293)

El objetivo de la posición original es proporcionar un medio de derivar principios de justicia que se abstraigan de lo contingente y por lo tanto de las influencias naturales y sociales moralmente irrelevantes (tal es la aspiración kantiana) sin tener que confiar en un dominio noumenal ni en la noción de un sujeto trascendente ubicado completamente más allá de la experiencia. La solución de Rawls consiste en restringir la descripción de las partes en la posición original a aquellas características que todos los seres humanos comparten en tanto seres racionales iguales y libres. En pocas palabras, éstas consisten en que cada uno es un ser que elige sus propios fines y que valora ciertos bienes primarios como instrumentos para su realización, sin importar cuáles puedan ser esos fines. Se asume que estos aspectos son comunes a todos los seres humanos en tanto tales, y en este sentido, no son contingentes.

> Así, dada la naturaleza humana, el querer estas cosas [bienes primarios] es una parte de su racionalidad [...] La preferencia por los bienes primarios se deriva, entonces, de las suposiciones más generales acerca de la racionalidad de las condiciones de la vida humana. Actuar a partir de los principios de la justicia es actuar a partir de imperativos categóricos en el sentido de que se aplican cualesquiera que sean sus propósitos en particular. Esto refleja el

hecho de que tales contingencias no aparecen como premisas en sus deducciones. (290-291)

Rawls reconoce que, a pesar de sus afinidades con Kant, la posición original se aparta de las perspectivas kantianas en diversos aspectos (293). Entre ellos seguramente debe aparecer la base que toma Rawls a partir de ciertas preferencias y deseos humanos generalizados para la derivación de los principios de justicia. Para Kant, fundar la ley moral sobre preferencias y deseos generalizados, no importa cuán difundidos puedan estar entre los seres humanos, sería meramente sustituir una heteronomía amplia por una más estrecha (Kant 1788: 55-58). No implicaría escapar de la contingencia en su sentido más exigente, que se aplica a la constitución de la naturaleza humana como también a la constitución de seres humanos particulares. Aun la «teoría débil del bien» sería demasiado robusta para satisfacer la concepción kantiana de la autonomía.

Esta contingencia más global no representa un problema para Rawls. Su interés es desarrollar una teoría de la justicia que sea imparcial entre las personas, y sólo necesita excluir aquellas contingencias que diferencian a las personas entre sí. Los atributos contingentes comunes a los seres humanos como tales no sólo no constituyen un problema para Rawls, sino que son elementos esenciales de su teoría moral. «La teoría moral debe tener la libertad de utilizar, a su gusto, supuestos contingentes y hechos generales.» Entre estos hechos generales están las circunstancias de la justicia en tanto hechos.

Aun cuando la perspectiva de Kant no las admita, la teoría de Rawls se basa en ellas; estos hechos aseguran que los principios de justicia que produce la teoría se apliquen a seres humanos en el mundo real, en lugar de aplicarse a seres trascendentes e incorpóreos más allá del mundo.

> Estando en las circunstancias de la justicia [las partes] están situadas en el mundo con otros hombres, quienes también se enfrentan a limitaciones tales como una escasez moderada o intereses competitivos. La libertad humana ha de ser regulada por principios elegidos a la luz de estas restricciones naturales. Así, la justicia como imparcialidad es una teoría de la justicia humana y entre sus premisas están los hechos elementales acerca de las personas y de su lugar en la naturaleza. (293-294)

De esta manera podemos ver por qué Rawls no puede adoptar simplemente una consideración kantiana de las circunstancias de la justicia que se acomode convenientemente con otras de sus posiciones kantianas; por qué está presionado a recurrir a la noción de una posición original que incluya como parte de su descripción una consideración empírica de las circunstancias humanas características. Es esta inquietante combinación la que hace surgir las objeciones que hemos considerado. En tanto que una concepción kantiana de la ley moral y del reino de los fines parece negar a la justicia su situación humana, por su parte la consideración humeana de la situación humana no parece capaz de dar cuenta de los fuertes supuestos a favor de la primacía de la justicia. Pero comprender cómo se producen estas inconsistencias no es disolverlas sino, quizás, confirmarlas. Y así parecería que las dos aspiraciones de la teoría de Rawls —evitar tanto la contingencia de los deseos existentes como la supuesta arbitrariedad y oscuridad de lo trascendente— después de todo no serían combinables: el punto de Arquímedes desaparecería en una letanía de contradicciones.

Las circunstancias de la justicia: la respuesta deontológica

A todo esto, Rawls podría contestar lo siguiente: las aparentes incompatibilidades entre la primacía de la justicia y sus circunstancias se basan en una mala interpretación de la posición original y del papel que ésta juega en la concepción como un todo. Las objeciones son demasiado apresuradas. No logran apreciar que la concepción de las circunstancias de la justicia se da dentro de la concepción de la posición original, la cual, debemos recordar, es desde el principio hipotética. Las condiciones descriptas en ella tienen el propósito de ser las condiciones en las que las partes en la posición original realizan sus deliberaciones, y no las condiciones reales en las que viven sus vidas los seres humanos comunes.

Las objeciones dependen en gran parte del supuesto erróneo de que los hechos de las circunstancias de la justicia tienen el propósito de ser hechos de la vida dentro del mundo

real y fenoménico en el cual los principios de la justicia se aplicarían realmente, y que su validez depende por lo tanto de las mismas consideraciones empíricas de las que dependen todos los supuestos fácticos normales. Pero la concepción de las circunstancias de la justicia no puede considerarse como una generalización empírica directa, que se establece o refuta a través de la evidencia de la sociología, la psicología, etc. Dado que la concepción general de las circunstancias de la justicia está ubicada dentro de la concepción de la posición original, las condiciones y las motivaciones que describe sólo pueden afirmarse sobre las partes de la posición original y no necesariamente sobre los seres humanos reales. Una vez instalada como una premisa de la posición original, la concepción de las circunstancias de la justicia deja de funcionar como una simple concepción empírica cuya exactitud puede verificarse con respecto a condiciones humanas reales. Su validez depende, por el contrario, en la medida en que la concepción de la cual es parte origine principios de justicia que capturen con éxito nuestras convicciones establecidas en un equilibrio reflexivo. «Lo mejor parece considerar estas condiciones simplemente como estipulaciones razonables, que han de ser valoradas, en su momento, por el conjunto de la teoría a la cual pertenecen... la justificación descansa en la concepción total y en la forma en que ésta se ajusta y organiza nuestros juicios en un equilibrio reflexivo» (578-9).

Por lo tanto, la concepción de las circunstancias de la justicia no necesita ser verdad en ningún sentido literal o empírico. La posición original que contiene es en todo caso una ficción admitida, un dispositivo heurístico diseñado para limitar nuestro razonamiento sobre la justicia en determinadas maneras. La distinción entre las provisiones de las circunstancias de la justicia y las motivaciones que realmente predominan en las sociedades humanas es para Rawls un tema constante: «Debemos tener presente que las partes en la posición original son individuos definidos teóricamente» (174). «La explicación de estas condiciones [es decir, de las circunstancias de la justicia] no implica ninguna teoría particular de la motivación humana» (155). La motivación de las personas en la posición original no debe confundirse con la motivación de las personas que en la vida diaria aceptan los principios de la justicia y que poseen el correspondiente sentido de la justicia» (175).

Rawls pone particular énfasis en que el supuesto del desinterés mutuo y de la ausencia de lazos morales anteriores en la posición original no implica un juicio acerca del desinterés mutuo o la falta de lazos morales entre las personas.

> Por supuesto que no es necesario suponer que en la vida cotidiana las personas no hacen sacrificios importantes unas por otras, ya que, por el contrario, lo hacen a menudo movidas por el afecto y por vínculos sentimentales. (207)

> No hay entonces ninguna inconsistencia en suponer que, una vez que se levante el velo de la ignorancia, las partes se den cuenta de que tienen vínculos sentimentales y afectivos, que desean promover los intereses de los demás y ver alcanzados sus fines. (154)

Aunque el supuesto del desinterés mutuo y la presunta preferencia por los bienes primarios son las premisas de motivación principales de la posición original, esto no implica que estas motivaciones se mantengan para las personas en la vida real, ni que sean aplicables a personas que vivan en una sociedad bien ordenada regida por los dos principios de la justicia.

> Con respecto al alcance del supuesto de la motivación, debemos tener en mente que *sólo se aplica a las partes en la posición original;* deben deliberar como si prefirieran más y no menos bienes primarios [...] *El supuesto puede no caracterizar la motivación general de las personas en sociedad, sin embargo, y en especial puede no aplicarse a los ciudadanos de una sociedad bien ordenada* (una sociedad regulada eficazmente por los principios públicos adoptados en la posición original). [Los subrayados son míos.] (Rawls 1975:543-4)

Esta podría ser la réplica de Rawls, con considerable beneficio. Renunciar a la lectura empirista directa de las circunstancias de la justicia equivale a rescatar el presupuesto de la primacía de la justicia al menos de las más obvias de las objeciones empiristas. También arrojaría cierta luz sobre expresiones tan enigmáticas como «*una* sociedad humana *está* caracterizada por las circunstancias de la justicia» [Los subrayados son míos] (155), que en el contexto de la concepción de Rawls parece más que una generalización meramente empírica y sin embargo no llega a ser una definición dispositi-

va. Pero el rechazo de la interpretación empirista lleva a una pregunta más difícil: si las premisas descriptivas de la posición original no están sujetas en ningún sentido directo a tests empíricos, ¿a qué tipo de tests están sujetas? Si los límites de los supuestos motivadores no son límites empíricos, ¿qué tipo de limitaciones son?

Todo lo que hemos podido afirmar hasta ahora sobre los fundamentos de una premisa de la posición original se reduce a que su validez depende de si la concepción de la cual forma parte origina principios de justicia que capturan con éxito nuestros juicios considerados en equilibrio reflexivo, o en qué medida lo hace. Pero esto no nos dice lo suficiente. Lo que evita que el método del equilibrio reflexivo se haga circular es la disponibilidad de criterios independientes de juicio en cada extremo, sin importar cuán provisorios sean éstos, a la luz de los cuales ajustamos y corregimos los otros. En el caso de la justicia, esto significa que debemos disponer de cierta manera (aunque sea provisoria) de juzgar, *tanto* lo deseable de los principios de la justicia que pueden resultar de una descripción particular, *como* lo plausibles o razonables que sean los supuestos motivadores que los generan. («Cada una de las suposiciones deberá ser por sí misma natural y plausible» – 35/36). Los criterios independientes, aunque provisorios, acerca de lo deseable de los principios resultan dados por nuestras intuiciones de lo que es justo. ¿Pero cuál es el fundamento correspondiente del aspecto descriptivo? Lo que buscamos es saber *con referencia a qué* se puede determinar lo plausible de los supuestos de la posición original. Nos vemos tentados a afirmar, de acuerdo con el aspecto normativo, que los criterios de lo plausible están dados por nuestras «intuiciones» sobre lo que es empíricamente verdadero. Pero como hemos visto, la tentación empirista de pensar que son las condiciones y motivaciones reales de los seres humanos las que proveen el estándar de lo plausible, da como resultado consecuencias inaceptables.

Para plantear nuestro problema de otra manera: como resulta claro de la concepción del equilibrio reflexivo, las condiciones de la posición original no pueden ser *tan* inmunes a las circunstancias humanas reales como para que aceptemos *cualquier* presupuesto que produzca principios de justicia atractivos. A menos que las premisas de tales principios tengan *cierta* semejanza con la condición de criaturas discerniblemente humanas, el éxito del equilibrio resulta dañado en esa medida.

Si pudiéramos ajustarnos a nuestras convicciones sobre lo justo solamente apelando a premisas que nos parecen excéntricas o extravagantes o metafísicamente extremas, nos sentiríamos justamente inclinados a cuestionar las convicciones a las que se adecuan tales principios. Después de todo esto se parece bastante a la crítica de Rawls a Kant, en el sentido de que era capaz de establecer conclusiones moralmente necesarias sólo al precio de una concepción de la circunstancia moral que tenía muy poco que ver con algo discerniblemente humano.

Resumiendo: la validez de una premisa de la posición original no está dada empíricamente, sino por un método de justificación conocido como «el equilibrio reflexivo». Este método implica dos maneras diferentes de justificación que se combinan para brindar corrección y respaldo mutuos. Un aspecto de la justificación apela a nuestras convicciones meditadas sobre la justicia; el otro apela a un estándar de lo plausible descriptivo pero no estrictamente empírico que estamos tratando de definir.

El mismo Rawls resulta algo ambiguo en su concepto del estándar descriptivo. Tanto en sus comentarios generales sobre la justificación como en su defensa de promesas específicas de la posición original, su lenguaje refleja esta confusión, y vale la pena un examen atento:

> Pero ¿cómo habremos de decidir cuál es la interpretación más favorable [de la situación inicial]? Supongo, entre otras cosas, que hay una amplia medida de acuerdo acerca de que los principios de la justicia habrán de escogerse bajo ciertas condiciones. Para justificar una descripción particular de la situación inicial hay que demostrar que incorpora estas *suposiciones comúnmente compartidas*. Se argumentará partiendo de *premisas débiles, aunque ampliamente aceptadas*, para llegar a conclusiones más específicas. Cada una de las suposiciones deberá ser por sí misma *natural y plausible*; algunas de ellas pueden incluso parecer inocuas o triviales. [Los subrayados son míos.] (35-36)

> En la búsqueda de la descripción más favorecida de esta situación trabajamos desde los dos extremos. Empezamos por describirla de tal modo que represente *condiciones generalmente compartidas y preferentemente débiles*. Vemos entonces si estas condiciones son suficientemente fuertes como para producir un conjunto significativo de principios. Si no, buscamos ulteriores premisas *igualmente razonables*. [Los subrayados son míos.] (38)

En su defensa específica del supuesto del desinterés mutuo, Rawls apela a criterios similares:

> Sin embargo, el postulado del desinterés mutuo en la posición original se formula con objeto de asegurar que los principios de la justicia no dependan de imposiciones más fuertes. No hay que olvidar que la posición original pretende incorporar *condiciones ampliamente aceptadas que sean no obstante débiles*. Una concepción de la justicia no debería, por tanto, suponer extensos vínculos de sentimientos naturales. Se trata de hacer *la menor cantidad posible de suposiciones* en la base de la teoría. [Los subrayados son míos.] (155)

Al argumentar a favor del desinterés mutuo frente a la benevolencia como la premisa de motivación adecuada, Rawls afirma que los supuestos combinados del desinterés mutuo y del velo de la ignorancia tienen «las ventajas de la simplicidad y la claridad» y al mismo tiempo aseguran los aspectos benéficos de motivaciones aparentemente más generosas. Si nos preguntamos por qué no se postula la benevolencia con el velo de la ignorancia, «la respuesta es que no hay necesidad de una *condición tan fuerte.* Lo que es más, desbarataría el propósito de fundamentar la teoría de la justicia en *estipulaciones débiles,* al mismo tiempo que resulta incongruente con las circunstancias de la justicia» [*Los subrayados son míos*] (176). Finalmente, Rawls afirma en sus conclusiones acerca de la justificación: «He señalado varias veces *la naturaleza mínima de las condiciones* en relación con los principios, cuando se consideran separadamente. Por ejemplo, el supuesto de una motivación mutuamente desinteresada *no es una estipulación exigente.* No sólo nos permite basar la teoría en una noción razonablemente precisa de la elección racional, sino que *pide poco a los individuos*» [Los subrayados son míos] (644).

Rawls parece ofrecer dos respuestas básicas a la cuestión de cómo debe justificarse un supuesto de la posición original a partir de un punto de vista descriptivo, y ninguna de las dos nos lleva demasiado lejos. La primera respuesta es que debe ser ampliamente aceptado y comúnmente compartido, la segunda es que debe ser débil por oposición a un supuesto fuerte, y de ser posible, natural, razonable, inocuo e incluso trivial. Sin embargo no resulta claro a qué apuntan estas consideraciones, o cómo en todo caso nos ayudan a determinar si debemos describir a las partes como mutuamente desinteresadas o benévolas.

En primer lugar, no resulta claro qué aspecto del supuesto motivador debe ser comúnmente compartido o ampliamente aceptado, y por qué esto juega en su favor. ¿Debemos concentrarnos en el *motivo* que es comúnmente compartido (en cuyo caso tendríamos que generalizar sobre los motivos que tienen las personas)? ¿O en el que sea considerado más ampliamente como el motivo prevaleciente (en cuyo caso tendríamos que generalizar sobre las generalizaciones que hacen los seres humanos sobre los motivos de los demás)? ¿O en el que se acepta ampliamente como una condición adecuada sobre principios de justicia (en cuyo caso tendríamos que generalizar sobre la manera en que las personas interpretarán probablemente los requisitos de un común acuerdo que estamos luchando por interpretar)? Pero estas interpretaciones son, o bien empíricas o bien fuente de nuevas preguntas, o ambas: y en todo caso no son obviamente relevantes para la validez de una premisa tal como la de desinterés mutuo o la de benevolencia como una condición de la posición original.

El requisito de que el supuesto sea débil (en lugar de fuerte) plantea justamente el interrogante que queremos responder: ¿débil o fuerte con respecto a *qué*? Podemos decir que un supuesto es débil desde un punto de vista conceptual, y por lo tanto que es probable que sea inocuo o trivial o adecuado de otra manera, cuando para su validez depende de la validez de unas pocas (relativamente) proposiciones relacionadas, y cuando aquéllas sobre las que se basa son a su vez débiles y no controvertidas. Un supuesto es fuerte en estos términos cuando, para que resulte verdadero, muchas otras cosas, incluyendo temas controvertidos, tienen que ser también verdad. Pero seguramente los supuestos de desinterés mutuo y de benevolencia no pueden distinguirse sobre la base de que uno o el otro sean supuestos más fuertes o débiles en el sentido conceptual: ninguno se basa en una premisa que sea *conceptualmente* más controvertida o problemática que el otro.

Si ninguno de los dos es más o menos exigente conceptualmente que el otro, el sentido alternativo parecería referirse a probabilidades estadísticas. Cuando los economistas del estado de bienestar, por ejemplo, se refieren a supuestos motivadores como débiles o fuertes, apuntan a describir la probabilidad de que la motivación sea aplicable a una proporción relativamente grande de la población. Mucho del lenguaje de Rawls parece sugerir este uso general probabilístico. ¿A qué

otra cosa puede referirse cuando afirma que asumir que las partes son mutuamente desinteresadas es asumir «menos» que suponer que existen amplios lazos de sentimiento natural? ¿Cómo sabe que asumir el desinterés mutuo no es una estipulación exigente ya que «exige poco de las partes»? ¿Estamos asumiendo que nos inclinamos por naturaleza al egoísmo más que a la benevolencia? Quizás para algunos resulta mucho más exacto postular que actúan en forma egoísta que en forma benévola. ¿La manera en que los seres humanos están naturalmente inclinados a conducirse resulta entonces una cuestión de probabilidades estadísticas? ¿Y cómo podría siquiera formularse esta cuestión en forma lo suficientemente precisa como para tener una estimación razonable sin especificar el rango de situaciones involucradas? De cualquier manera, si lo que quiere decir Rawls con su distinción entre los supuestos débiles y fuertes es que las premisas deben ser más realistas en lugar de menos realistas, entonces nos encontramos nuevamente en una comprensión meramente empírica de las condiciones de la posición original, que ya hemos rechazado.

Aunque la propia concepción de Rawls acerca de la posición original y del rango de sus premisas descriptivas no resulta clara, es esencial cierta concepción de estos temas para poder lograr que la teoría tenga cualquier viso de coherencia. A menos que podamos superar los aparentes obstáculos a la explicación de la posición original y sus premisas, la coherencia de toda la concepción estará en duda. Lo que necesitamos es una concepción de cuáles son exactamente las restricciones a los supuestos descriptivos adecuados a la situación inicial, es decir, qué los restringe *además* de las restricciones impuestas desde lo normativo, que consisten en nuestras convicciones consideradas sobre la justicia. De manera más general, necesitamos saber algo más preciso sobre la situación de la posición original, ya sea fenoménica u otra; ¿en qué *consiste* la posición original, de todas maneras?

En busca del sujeto moral

Creo que estas cuestiones pueden responderse, si bien no en las palabras explícitas de Rawls al menos en términos co-

herentes con su concepción como un todo. Al encontrar las respuestas podemos enfrentar ciertas divergencias del texto a fin de lograr su sentido. La justificación de esta interpretación se encontrará en sugerencias y rastros de evidencia sembrados en todo el texto y, lo que es más importante, en el sentido de que nos permite comprender la teoría de Rawls como un todo, y en especial su capacidad de resolver ciertos rasgos problemáticos de la posición original que no hemos podido resolver de ninguna otra manera.

Pero nuestro punto de divergencia es uno que está firmemente ubicado dentro del texto, en la noción del equilibrio reflexivo como método de justificación que gobierna la concepción como un todo. La clave es percibir la posición original como el eje del equilibrio reflexivo, en la medida en que puede lograrse. La posición original es el punto de apoyo del proceso de justificación en la medida que constituye un mecanismo a través del cual debe pasar toda justificación, el lugar al que deben llegar todos los argumentos y desde donde deben partir. A esto se debe que una premisa de la posición original pueda defenderse o atacarse desde cualquiera de las dos direcciones, sobre la base de cuán plausible resulta (en un sentido aún no determinado), o sobre la base de su ajuste con nuestras convicciones consideradas sobre lo justo.

> En la búsqueda de la descripción más favorecida de esta situación trabajamos desde los dos extremos. [...] Yendo hacia atrás y hacia delante, unas veces alterando las condiciones de las circunstancias contractuales, y otras retirando nuestros juicios y conformándolos a los principios, supongo que eventualmente encontraremos una descripción de la situación inicial que a la vez exprese condiciones razonables, y produzca principios que correspondan a nuestros juicios debidamente conformados y adaptados. Me referiré a este estado de cosas como «equilibrio reflexivo». (38).

La descripción de la posición original es el producto de dos ingredientes básicos: nuestro mejor juicio de lo que es «razonable y plausible» (que permanece sin explicar) por una parte, y nuestras convicciones consideradas sobre la justicia por el otro. A partir de la materia prima de nuestras intuiciones, filtradas adecuadamente y modeladas por la posición original, emerge un producto final. Pero es un producto final de dimensiones duales, y en esto yace la clave de nuestra explicación,

dado que lo que surge en un extremo en una teoría de la justicia debe surgir por el otro en una teoría de la persona, o más precisamente, una teoría del sujeto moral. Desde una perspectiva a través del lente de la posición original podemos ver los dos principios de la justicia; desde la otra perspectiva vemos un reflejo de nosotros mismos. Si el método del equilibrio reflexivo opera con la simetría que Rawls le atribuye, la posición original debe producir no solamente una teoría moral sino también una antropología filosófica.

A través de la mayor parte de su libro, Rawls se concentra especialmente en este punto. Su objetivo es producir una teoría de la justicia, por lo cual la mayor parte de su atención la dedica al argumento desde la posición original a los principios de la justicia, y a una descripción de la posición original que anticipe adecuadamente los requisitos de la justicia. Es comprensible que esté menos preocupado con el desarrollo del argumento en la dirección opuesta, y por lo tanto es menos explícito sobre qué se encontraría en este caso. Esto puede explicar en parte su mayor claridad en la base de nuestras intuiciones morales que sobre el origen y el rango de nuestras intuiciones «descriptivas» (qué las hace razonables o poco razonables, fuertes o débiles, etc.) Si esta reconstrucción es correcta, la norma independiente aunque provisional por la cual se evalúa lo razonable de nuestro supuesto descriptivo resulta dada no por las leyes de la psicología o la sociología empíricas, sino por la naturaleza del sujeto moral tal como lo entendemos, lo que equivale a afirmar: por la comprensión constitutiva que tenemos de nosotros mismos.

Dado que el objetivo de esta concepción es producir una teoría de la justicia, existe una tendencia a desechar los supuestos motivadores a veces poco atractivos de la posición original como parte de un mecanismo meramente heurístico que no guarda interés independiente o permanente alguno una vez que se han producido completamente los principios de la justicia. Pero si el equilibrio reflexivo funciona verdaderamente en ambos sentidos, la explicación de la circunstancia humana que emerge una vez que se logra el equilibrio reflexivo no puede ser ya descartada como el producto incidental de una invención, más de lo que se podrían descartar los principios de la justicia mismos. Dada la simetría metodológica de la posición original, no podemos considerar que uno de sus productos es la paja y el otro el trigo, y que el primero puede desecharse

una vez que se ha molido la harina. Debemos estar preparados para vivir con la visión contenida en la posición original, incluido el desinterés mutuo, preparados para vivir con ella en el sentido de aceptar su descripción como un reflejo apropiado de la circunstancia moral humana, que es coherente con nuestra comprensión de nosotros mismos.

> Por último, podemos recordar que la hipotética naturaleza de la situación original invita a preguntar: ¿por qué hemos de tener algún interés en esto, ni moral ni de otro género? Recordemos la respuesta: *las condiciones incorporadas a la descripción de esta situación son unas condiciones que nosotros, realmente, aceptamos*. O, *si no las aceptamos, podemos persuadirnos de hacerlo*, mediante las consideraciones filosóficas del tipo ocasionalmente introducido. [Los subrayados son míos.] (648)

Las condiciones incorporadas en la posición original guardan la misma relación con respecto a la teoría de la justicia de Rawls que los conceptos de la razón pura especulativa tienen con la teoría moral de Kant. «[...] no habrán de considerarse [...], como suele ocurrir en una construcción precipitada, para poner posteriormente puntales y apoyos, sino como verdaderos miembros que dejan ver la conexión del sistema [...]» (Kant 1787: 94).

Estas consideraciones sugieren firmemente que en la teoría de la justicia de Rawls se encuentra implícita una concepción del sujeto moral que tanto modela los principios de la justicia como es modelado a su imagen a través del medio de la posición original. Es esta concepción la que me propongo iluminar y explorar. Si puede hacerse explícita de alguna manera, ayudaría no solamente a resolver las perplejidades relativas al rango de la posición original, sino también a evaluar los presupuestos centrales de la concepción como un todo. Mientras que el argumento principal de Rawls tiende a asumir la naturaleza del sujeto moral como dada, y a llegar a los principios de la justicia a través de la posición original, me propongo trabajar en el sentido opuesto: tomar los principios de la justicia como provisionalmente dados y volver sobre nuestros pasos hacia la naturaleza del sujeto moral. Al hacerlo, asumo que estoy volviendo sobre los lineamientos de un argumento del siguiente tipo: asumiendo que somos seres capaces de justicia, y más precisamente, seres para quienes la justicia es primaria,

debemos ser criaturas de un cierto tipo, relacionados con las circunstancias humanas en cierta forma. ¿Qué debe ser verdad acerca de un sujeto para quien la justicia es la primera virtud? ¿Y cómo se incorpora la concepción de este tipo de sujeto en la posición original?

Ahora bien, la descripción de este sujeto tendrá un rango lógico específico. En cierto sentido será necesario, no condicional, y anterior a cualquier experiencia particular —éste es el propósito de que se formule como un «debe ser»— pero no constituirá por supuesto una demanda analítica. En cierto sentido será empírico, pero no «meramente» empírico. Dado el carácter reflexivo de tales descripciones, no son meramente descriptivas sino que también constituyen parcialmente el tipo de seres que somos. Nuestro conocimiento acerca de ellas es parte de lo que las hace verdaderas, y lo que nos hace las criaturas capaces de reflexión y autoconocimiento que somos.

Podemos describir una explicación general de estos rasgos constitutivos de nuestro autoconocimiento con una diversidad de nombres: una teoría de la persona, una concepción del «yo», una epistemología moral, una teoría de la naturaleza humana, una teoría del sujeto moral, una antropología filosófica. Estas descripciones conllevan diversas connotaciones, a veces conflictivas, habitualmente asociadas con las tradiciones filosóficas de las cuales se deducen. El hablar de la naturaleza humana, por ejemplo, a menudo equivale a sugerir una concepción teleológica clásica, asociada con la noción de una esencia humana universal, invariable en todo tiempo y lugar. Al hablar del «yo», por otro lado, tendemos a inclinar el tema a favor de nociones individualistas y a sugerir que la autocomprensión implicada se refiere meramente a la toma de conciencia de una persona individual, como por ejemplo en la psicoterapia. Estas asociaciones plantean ciertas dificultades, ya que amenazan con evitar la pregunta específica que deseamos responder, v.g., cómo se constituye el sujeto, en cuáles términos y en qué escala se concibe adecuadamente. Para evitar la confusión a este respecto, debería decir anticipadamente que la explicación a la que me refiero constituye una antropología filosófica en el sentido amplio; filosófica, dado que se llega a ella a través de la reflexión y no por medio de la generalización empírica; antropología, porque se refiere a la naturaleza del sujeto humano en sus diversas formas posibles de identidad.

El yo y el otro: la prioridad de la pluralidad

Partiendo de estas calificaciones, y con referencia a nuestra argumentación anterior sobre el problema del «yo», podemos proceder a reconstruir el razonamiento de Rawls sobre la naturaleza del sujeto moral. Para Rawls, la primera característica de toda criatura capaz de justicia reside en que es plural en número. La justicia no podría aplicarse en un mundo donde existiera un solo sujeto. Sólo podría darse en una sociedad de seres que fueran de alguna manera distinguibles unos de otros. «Los principios de la justicia se ocupan de las pretensiones conflictivas producto de las ventajas obtenidas por la cooperación social; se aplican a las relaciones entre varias personas o grupos. La palabra "contrato" sugiere [...] esta pluralidad» (34). Para que haya justicia, debe existir la posibilidad de que se den pretensiones conflictivas, debe haber más de un sujeto demandante. De esta manera, la pluralidad de personas puede considerarse un presupuesto necesario para la posibilidad de justicia.

Rawls insiste sobre la pluralidad esencial del sujeto humano cuando acusa al utilitarismo de extender a la sociedad como un todo los principios de selección racional aplicables a un solo hombre. Esto es una falacia, alega, porque fusiona diversos sistemas de justicia en un único sistema de interés, y no logra tomar seriamente las diferencias entre las personas. Dentro del utilitarismo «[...] *muchas personas* son fundidas en una sola [...]», e «[...] *individuos aislados* son considerados como otras tantas líneas diferentes [...]». Pero el utilitarismo se equivoca, dado que: «No hay razón para suponer que los principios que debieran regular una asociación de hombres sean simplemente una extensión del principio de elección de un solo hombre. Por el contrario: si suponemos que el principio regulador correcto para cualquier cosa depende de *la naturaleza de la cosa*, y que *la pluralidad de personas distintas con sistemas de fines separados es una característica esencial de las sociedades humanas*, no deberíamos esperar que los principios de la elección social fuesen utilitaristas» [Los subrayados son míos] (45-57).

El tomar seriamente «la pluralidad y particularidad de los individuos» va más allá de defender la libertad e independencia de pensamiento, y de sostener que el bien de la sociedad consiste en las ventajas que disfrutan los individuos, como ha-

cían los utilitaristas. Significa comprender la pluralidad de personas como un rasgo esencial de cualquier explicación del sujeto moral, un postulado de la antropología filosófica. En este sentido, «el utilitarismo no es individualista, al menos cuando se llega a él a través del curso de reflexión más natural [...]», ya que al fusionar todos los sistemas de intereses en uno solo, se contradice su postulado esencial (48).

Pero para que los sujetos sean plurales, debe existir algo que los diferencie, alguna manera de distinguirlos unos de otros, algún principio de individualización. Para Rawls, nuestras características individualizadoras están empíricamente dadas, por la concatenación diferente de necesidades y deseos, objetivos y atributos, propósitos y fines que caracterizan a los seres humanos en su particularidad. Cada individuo se ubica de forma única en el tiempo y el espacio, nace en una familia y en una sociedad en particular, y las contingencias de estas circunstancias, junto con los intereses y valores y las aspiraciones a los que dan origen, son lo que diferencia a las personas, lo que las hace las personas individuales que son.

Dentro de cada grupo de personas, especialmente aquéllos en circunstancias similares, encontraremos probablemente ciertas características que se superponen, ciertos intereses que resultan comunes. Pero a pesar de la similitud de la situación, no importa cuán próxima sea ésta, no podemos decir que dos personas hayan estado nunca situadas en forma idéntica, ni que dos personas hayan tenido objetivos e intereses idénticos en todos los aspectos, porque de ser así, ya no resultaría claro cómo podríamos identificarlas como dos personas diferentes. De esta manera, la pluralidad esencial de las personas queda asegurada, o mejor, definida. Para Rawls, nuestra pluralidad fundamental es un presupuesto necesario del hecho que seamos criaturas capaces de justicia. En qué consiste la individualidad de cada persona en particular, sin embargo, es una cuestión empírica. El rasgo fundamental del sujeto moral es su pluralidad, y dados los medios de individualización, la cantidad de su pluralidad se corresponde con la cantidad de seres humanos empíricamente individualizados en el mundo. Todo esto equivale a afirmar que en la perspectiva de Rawls acerca del sujeto moral, cada ser humano individual es un sujeto moral, y cada sujeto moral es un ser humano individual.

En este punto puede resultar razonable preguntarnos si, dentro de la concepción de la persona de Rawls, la unidad pue-

de aparecer como una característica esencial del sujeto moral al igual que la pluralidad, igualmente necesaria como presupuesto dentro de la concepción del hombre como una criatura capaz de justicia. De hecho, Rawls describe las circunstancias de la justicia como «las condiciones normales bajo las cuales la cooperación humana resulta tanto posible como necesaria», y a partir de esto podría pensarse que cuando la pluralidad esencial del sujeto moral hace necesaria la cooperación humana, cierta unidad esencial de las personas hace que esta cooperación sea posible. Pero esto sería malinterpretar la lógica de la concepción de Rawls, y amenazaría con minar las prioridades de las cuales depende la ética deontológica. Creo que la respuesta que daría sería aproximadamente la siguiente.

Aunque es verdad que el principio de unidad ocupa un lugar importante en la justicia como imparcialidad (véase en particular la idea de la unión social, en la Sección 79), resultaría erróneo otorgarle una prioridad equivalente a la de la pluralidad; no resultan esenciales para nuestra naturaleza de la misma manera. Esto se debe a que cualquier fundamentación de la unidad de la subjetividad humana debe *presuponer* su pluralidad, mientras que no se verifica lo contrario. Esto puede comprobarse si consideramos la noción de una sociedad humana como una empresa cooperativa para beneficio mutuo, señalada típicamente por el conflicto al igual que por la identidad de intereses. El conflicto de intereses surge, como hemos visto, a partir del hecho de que los sujetos de la cooperación tienen diferentes intereses y fines, y *este* hecho surge de la naturaleza de un ser que tiene la capacidad de la justicia. La identidad de intereses, sin embargo, expresa el hecho de que las partes tienen necesidades e intereses adecuadamente similares, de forma tal que la cooperación entre ellos es mutuamente beneficiosa. Y *este* hecho, el que sus necesidades e intereses coincidan de esta manera, no se deriva de la naturaleza de su subjetividad, sino simplemente de un accidente feliz de sus circunstancias. Su capacidad de reunirse para cooperar en aras de mutuos beneficios presupone una pluralidad antecedente. La cooperación es en su misma naturaleza una cooperación entre agentes, cuya pluralidad debe por lo tanto ser anterior a la identidad de intereses que se realizan a través de la asociación cooperativa.

> La idea fundamental es la de que queremos estimar los valores sociales del bien intrínseco en las actividades institucionales, co-

munitarias y asociativas mediante una concepción de la justicia que en sus bases teóricas es individualista. Por razones de claridad, entre otras, no queremos basarnos en un concepto no definido de la comunidad, o suponer que una sociedad es un todo orgánico con una vida propia, distinta y superior a la de sus miembros en sus relaciones con los demás. Así, la concepción contractual de la posición original es considerada en primer lugar. [...] Partiendo de esta concepción, aunque pueda parecer individualista, podemos explicar eventualmente el valor de la comunidad. (301-302)

El hecho de que somos personas distintas, caracterizadas por sistemas de fines diferentes, constituye un presupuesto necesario de un ser capaz de justicia. En qué consisten específicamente nuestros fines, y si aciertan a coincidir o superponerse con los fines de otros, son cuestiones empíricas que no pueden saberse por anticipado. Es en este sentido —más epistemológico que psicológico— que la pluralidad de sujetos resulta dada con prioridad a su unidad. Somos en primer lugar individuos distintos, y luego (si las circunstancias así lo permiten) formamos relaciones y nos involucramos en empresas de cooperación con otros. El punto no es que las personas cooperen solamente debido a motivos egoístas, sino que nuestro conocimiento sobre la base de la pluralidad está dado como previo a la experiencia, mientras que nuestro conocimiento sobre la base de la unidad o la cooperación sólo puede surgir a la luz de ésta. En cualquier caso particular, simplemente debemos comprobar si existe o no la base para la cooperación.

El yo y sus fines: el sujeto de la posesión

En la ética deontológica «el "yo" es anterior a los fines que mediante él se afirman» (619). Para Rawls, brindar una fundamentación de esta prioridad plantea un particular desafío, ya que su proyecto elimina la posibilidad del «yo» que logra su prioridad al habitar un dominio trascendente o noumenal. En esta perspectiva, cualquier explicación del «yo» y sus fines debe aclararnos no una sino dos cosas; cómo se distingue el «yo»de sus fines, y también cómo se conecta el «yo» con sus fines. Sin

la primera, sólo nos resta un sujeto radicalmente situado; sin la segunda, un sujeto radicalmente no corporeizado.

La solución de Rawls, implícita en el diseño de la posición original, consiste en concebir al «yo» como el sujeto de la posesión, ya que en la posesión el «yo» se distancia de sus fines sin estar completamente separado. La noción del «yo» como el sujeto de la posesión puede ubicarse en el supuesto del desinterés mutuo. Este supuesto se parece superficialmente a un presupuesto psicológico —estipula que las partes no tienen interés en los intereses de las demás— pero dada su ubicación dentro de la posición original funciona sin embargo como un supuesto epistemológico, un supuesto sobre las formas de autoconocimiento de las que somos capaces. Es así como Rawls puede mantener de manera coherente que el supuesto del desinterés mutuo es «la condición de motivación principal de la posición original» (189), y sin embargo «no implica ninguna teoría particular de la motivación humana» (155).

Podemos observar cómo se da esto. El supuesto del desinterés mutuo no se formula acerca de lo que motiva a las personas, sino que es un supuesto acerca de la naturaleza de los sujetos que poseen motivaciones en general. Se ocupa de la naturaleza del «yo» (es decir, la forma en que está constituido, cuál es su postura con respecto a la situación en general), y no de la naturaleza de los deseos u objetivos del «yo». Se refiere al *sujeto* de los intereses y fines, y no al *contenido* de éstos, cualesquiera sean. Mientras Kant argumenta que toda experiencia debe ser la experiencia de algún sujeto, el supuesto de Rawls sobre el desinterés mutuo sostiene que todos los intereses deben ser los de algún sujeto.

> Más aún, aunque no se supone que los intereses propuestos por estos planes sean intereses *en sí mismos*, son intereses *de un «yo»* que considera su concepción del bien como digna de reconocimiento [...] (153)
>
> Nótese que no he hecho ningún supuesto restrictivo acerca de las concepciones del bien que tienen las partes, excepto que son planes racionales a largo plazo. Tales planes determinan los objetivos e intereses *de un «yo»*, aunque no se presume que éstos sean egoístas. Que sea o no así depende del tipo de fines que una persona persiga. Si los propósitos finales de una persona son la riqueza, la posición, la influencia y el reconocimiento del prestigio social, seguramente su concepción del bien será egoísta. Sus intereses dominantes están *en sí mismos*, y no son meramente, *como*

debieran, los intereses *de un «yo»* [Los subrayados son míos.] (154-155)

En el supuesto del desinterés mutuo, encontramos la clave de la concepción del sujeto de Rawls, una visión de la manera en que debemos ser a fin de constituirnos en sujetos para quienes la justicia es primaria. Pero la noción del «yo» como sujeto de la posesión, tomada aisladamente, no completa este panorama. Como sugiere la explicación de la pluralidad, no servirá cualquier sujeto de la posesión, sino solamente un sujeto con una individualidad antecedente, cuyos límites del «yo» estén fijados antes de la experiencia. Para constituir un «yo» deontológico, debo ser un sujeto cuya identidad esté dada independientemente de las cosas que tengo, es decir, independientemente de mis intereses y fines, y de mis relaciones con los demás. Combinada con la idea de posesión, esta noción de la individualidad completa poderosamente la teoría de Rawls sobre la persona. Podemos apreciar plenamente sus consecuencias al comparar los dos aspectos de la posesión —dos maneras distintas de que un interés sea «de un yo»— y observar cómo la noción de la individualidad antecedente compromete al «yo» deontológico con uno de ellos.

En tanto que poseo algo, estoy al mismo tiempo relacionado con él y distanciado de él. Decir que poseo una cierta inclinación, deseo o ambición es decir que estoy relacionado con él de cierta manera —es *mío* por oposición a *tuyo*— y también que estoy distanciado de él de cierta manera —es *mío* en lugar de ser *yo mismo*. Este último punto significa que si pierdo algo que poseo, sigo siendo el mismo «yo» que lo tenía; éste es el sentido, al principio paradójico pero que tras la reflexión surge inevitablemente, en el que la noción de posesión es una noción de distanciamiento. Este aspecto de distancia es esencial para la continuidad del «yo», y preserva una cierta dignidad e integridad del «yo» al impedir su variación frente a la menor contingencia. Preservar esta distancia, y la integridad que implica, requiere típicamente un cierto tipo de autoconocimiento. Para preservar la distinción entre lo que soy *yo* y lo que es (solamente) *mío*, debo saber algo acerca de quién soy, o ser capaz de distinguirlo cuando la ocasión lo exija. Ulises pudo sobrevivir a su peligroso viaje de regreso a su hogar utilizando varios disfraces, y su habilidad para hacerlo presuponía una comprensión de quién era él. Ya que este autoconocimiento precede a su experiencia en

este sentido, pudo regresar a su ciudad siendo la misma persona que había partido, ser reconocido por Penélope, inalterado por su viaje, a diferencia de Agamenón, que se volvió un extraño para su familia y encontró una suerte diferente.[5]

Una consecuencia del aspecto dual de la posesión es que puede disiparse o disminuir de dos maneras diferentes. Pierdo gradualmente posesión de una cosa no solamente en la medida en que se distancia de mi persona, sino también a medida que la distancia entre mi «yo» y la cosa se reduce y ambos tienden a la colisión. Pierdo posesión de un deseo o una ambición a medida que se desvanece mi compromiso con éste, a medida que se debilita mi control, pero también, al traspasar un cierto punto, a medida que crece mi vinculación con éste, a medida que gradualmente se adiciona a mí. A medida que un deseo o ambición forma cada vez más parte de mi identidad, se vuelve cada vez más *yo* y cada vez menos *mío.* O, como puede decirse en algunos casos, gradualmente no soy yo el que lo posee, sino que yo soy poseído *por* él. Imaginemos que un deseo, que se tiene tentativamente al principio, se vuelve paulatinamente más central para mis objetivos generales, hasta que finalmente se transforma en la consideración más importante en todo lo que pienso o hago. A medida que crece como deseo y se transforma en una obsesión, yo lo poseo cada vez menos y él me posee cada vez más, hasta que finalmente se vuelve indistinguible de mi identidad.

Veamos otro tipo de ejemplo: en la medida en que la Declaración de la Independencia de Estados Unidos es correcta, en cuanto a que el hombre ha recibido de su Creador ciertos derechos inalienables entre los cuales se encuentran la vida, la libertad, y la búsqueda de la felicidad, esta enunciación describe no lo que *tenemos* como hombres libres sino lo que *somos.* Un don no es tanto una posesión como una *naturaleza* de cierto tipo; aquél que reniegue de su libertad o se aboque a la búsqueda de una existencia desdichada experimentará estos dones como restricciones más que como posesiones. En la medida en que estos derechos son verdaderamente inalienables, un hombre no tiene más derecho a deshacerse de ellos en su caso particular que a privar de ellos a los demás. El suicidio es comparable al asesinato, y venderse a uno mismo como esclavo resulta el equivalente moral a esclavizar a otro.

5. Debo agradecer a Allen Grossman por este ejemplo.

Como sugieren estas imágenes, la posesión está limitada por la facultad de agencia humana y por un sentido de autocontrol. La pérdida de la posesión, desde ambos puntos de vista, puede entenderse como una forma de pérdida de poder. Cuando se desvanece mi posesión de un objeto, ya sea porque se escurre fuera de mi alcance o porque me veo superado por su magnitud, sin poder ante él, mi facultad de agencia con respecto a ese objeto también se ve disminuida. Cada desafío está asociado con una noción diferente de la agencia, que implica, a su vez, una distinta fundamentación de la relación del «yo» con sus fines. Podemos pensar en las dos dimensiones de la agencia como maneras diferentes de impedir el deslizamiento hacia la pérdida de posesión, y distinguirlas por la manera en que funcionan para restablecer un sentido de autocontrol.

El primer tipo de pérdida de posesión implica que el fin se distancia del «yo» cuyo fin fue alguna vez. Se vuelve cada vez más difuso en el sentido de que es mi fin y no el tuyo, o el de otra persona, o el de nadie. El «yo» pierde poder porque se ve disociado de esos fines y deseos que se entretejen para formar un todo coherente y que le brindan una firmeza de propósitos, forman un plan de vida, y por lo tanto justifican la continuidad del «yo» con sus fines. Mientras que el «yo» es considerado como dado previamente a sus fines, con sus límites fijados permanentemente de modo que sean impermeables, invulnerables a la transformación por la experiencia, esta continuidad es perpetua e inherentemente problemática; la única manera en que puede afirmarse es a través del «yo», para llegar más allá de sí mismo, para tomar como un objeto de su voluntad los fines que desea poseer, y mantenerlos como externos a sí, tal como debe ser.[6]

El segundo tipo de pérdida de posesión opera como una privación de poder de otra manera. En este caso el problema no consiste en superar la distancia creada por el distanciamiento entre el fin y el «yo», sino en recuperar y preservar un espacio que amenaza cada vez más con desaparecer. Superado por los intereses y presiones de distintos propósitos y fines posibles, todos los cuales presionan indiscriminadamente sobre mi iden-

6. Compárese con Kant (1797: 62) «Por lo tanto, la relación de tener algo externo a uno mismo como nuestra (propiedad) consiste en una unión puramente *de jure* de la voluntad del sujeto con ese objeto, independientemente de su relación con éste en el tiempo y el espacio, y de acuerdo con el concepto de posesión inteligible».

tidad, soy incapaz de distinguirlos u ordenarlos, incapaz de marcar los límites o fronteras de mi «yo», incapaz de determinar dónde termina mi identidad y dónde comienza el mundo de los atributos, los objetivos y los deseos. He perdido poder en el sentido de que me falta una clara noción de quién soy yo específicamente. Demasiadas cosas son demasiado esenciales para mi identidad. Cuando los fines están dados con prioridad al «yo» que constituyen, los límites del sujeto se encuentran abiertos, su identidad es infinitamente moldeable y extremadamente fluida. Incapaz de distinguir lo que es mío de lo que soy yo, estoy en constante riesgo de ahogarme en un mar de circunstancias.

Podemos comprender la facultad de agencia humana como la facultad por la cual el «yo» realiza sus fines. Esto reconoce su estrecha conexión con la noción de posesión sin eludir la cuestión de cuál dimensión de la posesión está en juego, ni de la prioridad relativa del «yo» y de los fines. Puesto que soy un ser con fines, existen al menos dos maneras de «realizarlos»: una es por elección, la otra por descubrimiento, al «revelarlos». El primer sentido de «realizar» es el que podemos llamar la dimensión voluntarista de la agencia, el segundo es la dimensión cognitiva. Cada tipo de función de agencia puede verse como reparadora de un tipo distinto de pérdida de posesión.

Cuando el «yo» pierde poder debido a que se aleja de sus fines, la pérdida de posesión se repara por medio de la facultad de agencia en su sentido voluntarista, en el cual el «yo» se relaciona con sus fines en tanto sujeto provisto de voluntad con los objetos de su elección. La facultad relevante involucra el ejercicio de la voluntad, ya que es la voluntad la que puede trascender el espacio entre el sujeto y su objeto sin requerir que éste sea cerrado.

Cuando el «yo» sufre una pérdida de poder debido a que se asimila con sus fines, la pérdida de posesión se repara por medio de la facultad de agencia en su sentido cognitivo, en el cual el «yo» se relaciona con sus fines como un sujeto que conoce los objetos de su entendimiento. Cuando los fines del «yo» están dados por adelantado, la función de agencia relevante no es la voluntarista sino la cognitiva, ya que el sujeto logra el autocontrol, no a través de la elección de lo que ya está dado (que sería ininteligible), sino reflexionando e investigando su naturaleza constitucional, discerniendo sus leyes e imperativos, y reconociendo sus propósitos como propios. Cuando la facultad

de la voluntad intenta revertir la separación entre el «yo» y los fines reconstruyendo una cierta continuidad entre ellos, la reflexión es una facultad distanciadora, y aparece en un determinado apartamiento. Funciona al restablecer el espacio perdido entre el «yo» y sus fines. En la reflexión, el «yo» vuelve su foco de atención introspectivamente sobre sí mismo, haciendo del «yo» su propio objeto de investigación y reflexión. Cuando puedo reflexionar sobre mi obsesión, puedo identificarla y hacerla el objeto de mi reflexión, soy por lo tanto capaz de establecer un cierto espacio entre ella y yo, y por lo tanto de disminuir su control. Se vuelve más un atributo y menos un constituyente de mi identidad, y también se disuelve transformándose, de una obsesión, en un mero deseo.

Cuando el sujeto se considera como anterior a sus fines, el autoconocimiento no es una posibilidad en este sentido, ya que los límites que definiría ya están tomados como dados, sin reflexión, por el principio de la individualización antecedente. Los límites del «yo» son fijos y dentro de ellos todo es transparente. La cuestión moral relevante no es «¿Quién soy yo?» (ya que la respuesta a esta pregunta está dada con anterioridad), sino «¿Qué fines elegiré?», y ésta es una pregunta dirigida a la voluntad.

Para un «yo» cuya identidad se constituye a la luz de los fines que ya están frente a él, la facultad de la agencia consiste no tanto en ejercer la voluntad como en lograr la comprensión de sí mismo. La pregunta relevante no es cuáles fines elegir, ya que mi problema es precisamente que la respuesta a esta pregunta ya está dada, sino quién soy yo, cómo puedo discernir dentro del conglomerado de fines posibles lo que soy yo de lo que es mío. En este caso los límites del «yo» no son fijos, sino posibles, sus contornos no son ya evidentes por sí mismos sino al menos parcialmente no formados. El volverlos evidentes y el definir los límites de mi identidad son una sola cosa. El autocontrol que en el primer caso se mide en términos del alcance de mi voluntad, en el segundo caso se determina por la profundidad y claridad de mi conocimiento de mí mismo.

Podemos ver ahora que el conglomerado de supuestos asociados con la noción voluntarista de la agencia y el aspecto distanciador de la posesión completan la teoría de la persona de Rawls. La noción de un sujeto de la posesión, individualizado anticipadamente y dado con prioridad a sus fines, parece exactamente la concepción necesaria para redimir la ética

deontológica sin caer en la trascendencia. De esta manera, el «yo» se distingue de sus fines —ocupa un lugar más allá de ellos, a la distancia, con una cierta prioridad— pero también está relacionado con los fines, como sujeto poseedor de una voluntad con respecto a los objetos de su elección.

La noción voluntarista de la agencia es por lo tanto el elemento clave de la concepción de Rawls, y juega un papel central en la ética deontológica como un todo. «No es nuestro propósito el que revela principalmente nuestra naturaleza [...]» (619), sino nuestra capacidad de elegir nuestros objetivos la que resulta más importante, y esta capacidad encuentra su expresión en los principios de la justicia. «Así, una persona moral es un sujeto con fines que él ha elegido, y su preferencia fundamental se inclina a favor de las condiciones que le permitan construir un modo de vida que exprese su naturaleza de ente racional, libre e igual, tan plenamente como las circunstancias lo consientan» (621). Ésta, finalmente, es la razón por la que no podemos considerar a la justicia simplemente como un valor entre otros. «Para realizar nuestra naturaleza, no tenemos más alternativa que la de proteger nuestro sentido de la justicia para que rija nuestros restantes objetivos» (634).

El individualismo y las demandas de la comunidad

En nuestra reconstrucción del sujeto deontológico descubrimos finalmente el estándar con respecto al cual pueden evaluarse las premisas descriptivas de la posición original, el contrapeso de nuestras instituciones morales que suministra una prueba para ambos extremos en el caso del equilibrio reflexivo de Rawls. Esta concepción del sujeto es la que refleja el supuesto del desinterés mutuo, y no una justificación de las motivaciones humanas en particular.

No olvidemos que en la explicación de Rawls, «el postulado del desinterés mutuo en la posición original se formula con objeto de asegurar que los principios de la justicia no dependan de imposiciones más fuertes» (155), y la finalidad de evitar los supuestos fuertes es la de posibilitar la derivación de principios que no presupongan ninguna concepción particular del

bien. «La libertad para adoptar una concepción del bien está limitada únicamente por principios deducidos de una doctrina que no impone restricciones previas sobre estos conceptos. La presunción de un desinterés mutuo en la posición original implica esta idea» (290). Los supuestos fuertes o controvertidos amenazarían con la imposición de una concepción particular del bien, inclinando por anticipado la elección de los principios.

¿Cuán débiles o fuertes *son* entonces los supuestos que forman la concepción de la persona de Rawls? ¿Con qué abanico de valores y fines resultan compatibles? ¿Son lo suficientemente débiles e inocentes como para evitar descartar por anticipado cualquier concepción del bien? Ya hemos visto que la lectura empirista de la posición original produce una sucesión de objeciones a este alcance; se alega que las circunstancias de la justicia, y especialmente el supuesto del desinterés mutuo, introducen un sesgo individualista, y que rechazan o desvirtúan de alguna manera motivos tales como la benevolencia, el altruismo y los sentimientos comunitarios. Tal como escribió un crítico, la posición original contiene «un fuerte sesgo individualista, que se ve reforzado por el supuesto de motivación del desinterés mutuo y la ausencia de envidia. [...] La posición original parece presuponer no solamente una teoría neutral del bien, sino una concepción liberal e individualista de acuerdo con la cual lo mejor que podemos esperar para una persona es que pueda seguir su propio camino sin restricciones, siempre que no interfiera con los derechos de los demás» (Nagel 1973: 9-10).

Sin embargo, como insiste acertadamente Rawls, su teoría no es la «doctrina estrechamente individualista» que la crítica empirista supone: «Una vez que se comprende el motivo del supuesto del mutuo desinterés, la objeción parece fuera de lugar» (644). A pesar de su dimensión individualista, la justicia como imparcialidad no defiende a la sociedad privada como un ideal (522), ni presupone motivaciones egoístas (155), o se opone a valores comunitarios. «Aunque la justicia como imparcialidad se inicia tomando a las personas en la posición original como individuos [...] esto no es obstáculo para explicar los sentimientos morales de orden más elevado que sirven para mantener unida a una comunidad de personas» (223).

Rawls ha enfatizado particularmente que el supuesto del desinterés mutuo no implica, con respecto a la elección de principios, un sesgo a favor de los valores individualistas a expen-

sas de los comunitarios. Aquéllos que suponen esto no tienen en cuenta el rango especial de la posición original, y asumen erróneamente que los motivos que se atribuyen a las partes se suponen de aplicación general a seres humanos reales o a personas en una sociedad bien ordenada. Pero no se trata de ninguno de estos casos. Los motivos que se atribuyen a las partes en la posición original no reflejan las motivaciones reales que se dan en la sociedad, ni determinan directamente los motivos de las personas en una sociedad bien ordenada.

Dado el alcance restringido de estos supuestos, afirma Rawls, «parece no haber ninguna razón directa por la que los fines de las personas en una sociedad bien ordenada debieran ser predominantemente individualistas» (1975: 544). Probablemente los valores comunitarios existirían, al igual que todos los demás valores que los individuos pueden decidir adoptar, e incluso posiblemente florecerían en una sociedad gobernada por los dos principios de la justicia.

> No existe razón alguna por la que una sociedad bien ordenada deba alentar primordialmente valores individualistas si esto significa formas de vida que lleven a los individuos a perseguir su propio camino y no preocuparse por los intereses de los demás (aunque respetando sus derechos y libertades). Normalmente se puede esperar que la mayoría de las personas pertenezcan a una o más asociaciones y que tengan al menos algunos fines colectivos en este sentido. Las libertades básicas no tienen el propósito de mantener a las personas aisladas unas de las otras, ni de persuadirlas a que lleven vidas privadas, aunque algunos individuos sin duda lo harán, sino el de asegurar el derecho de libre circulación entre las asociaciones y las comunidades más pequeñas. (1975: 550).

En la concepción de la persona de Rawls, mis fines son benévolos o comunitarios cuando toman como objeto el bien de otro, o de un grupo de otros con los cuales puedo estar asociado, y de hecho no hay nada en su concepción que descarte los fines comunitarios en este sentido. Todos los intereses, valores y concepciones del bien están abiertos al «yo», en tanto puedan ser clasificados como los intereses de un sujeto anticipadamente individualizado y estén dados con prioridad a sus fines; v.g., en tanto describan los objetos que busco y no el sujeto que soy. Solamente los *límites* del «yo» están anticipadamente fijados.

No obstante, esto sugiere un sentido más profundo en el cual la concepción de Rawls es individualista. Podemos ubicar este individualismo e identificar las concepciones del bien que se excluyen al recordar que el «yo», de acuerdo con Rawls, no es solamente un sujeto de la posesión, sino un sujeto anticipadamente individualizado, que permanece siempre a cierta distancia de los intereses que posee. Una consecuencia de esta distancia es que pone al «yo» más allá del alcance de la experiencia; lo hace invulnerable, fija su identidad definitivamente. Ningún compromiso hará mella como para que no pueda comprenderme a mí mismo en su ausencia. Ningún cambio de los propósitos y planes de vida puede ser tan perturbador que altere los contornos de mi identidad. Ningún proyecto puede ser tan esencial que el abandonarlo amenace a la persona que soy yo. Dada mi independencia de los valores que poseo, siempre puedo distanciarme de ellos; mi identidad pública como una persona moral «no se ve afectada por los cambios que con el tiempo se operan en mi concepción del bien» (Rawls 1980: 544-5).[7]

Pero una noción del «yo» tan independiente como ésta descarta toda concepción del bien (o del mal) limitada por la posesión en el sentido constitutivo. Elimina la posibilidad de cualquier lazo (u obsesión) capaz de llegar más allá de nuestros valores y sentimientos para comprometer nuestra misma identidad. Elimina la posibilidad de una vida pública en la cual, para bien o para mal, estén en juego tanto la identidad como los intereses de los participantes. Y excluye la posibilidad de que los propósitos y fines comunes puedan inspirar una comprensión de sí más o menos expansiva, definiendo así una comunidad en el sentido constitutivo, una comunidad que describe al sujeto y no simplemente a los objetos de las aspiraciones compartidas. En forma más general, podemos afirmar que la justificación de Rawls descarta la posibilidad de lo que podríamos llamar formas de autocomprensión «intersubjetivas» o «intrasubjetivas», maneras de concebir al sujeto que no suponen que sus límites están dados con anticipación. A diferencia de Rawls, las concepciones intersubjetivas e intrasubjetivas no dan por supuesto que hablar del «yo», desde un

7. Rawls sugiere en cierto momento que mi identidad privada en tanto persona moral podría no ser inmune a los lazos constitutivos de la misma manera (1980: 545). Véase página 182.

punto de vista moral, equivalga necesaria e inevitablemente a hablar de un «yo» individualizado anticipadamente.

Las concepciones intersubjetivas permiten que en ciertas circunstancias morales, la descripción relevante del «yo» abarque más de un solo ser humano individual, como en el caso en que atribuye responsabilidad o afirma la existencia de una obligación hacia una familia, comunidad, clase o nación en lugar de hacia un ser humano en particular. Tales concepciones son presumiblemente lo que Rawls tenía en mente cuando rechaza, «por razones de claridad, entre otras», lo que denomina «un concepto no definido de la comunidad [...]» y la noción de que «una sociedad es un todo orgánico» (301-302), ya que esto sugiere el aspecto metafísicamente perturbador de Kant que Rawls se consagra a reemplazar.

Las concepciones intrasubjetivas, por otra parte, aceptan que para ciertos propósitos la descripción adecuada del sujeto moral puede referirse a una pluralidad de distintos «yoes» dentro de un ser humano individual y único, de la misma manera en que explicamos el debate introspectivo en términos de una puja entre identidades competitivas, o en que describimos los momentos de introspección como un autoconocimiento obstruido, o en que absolvemos a una persona de responsabilidad por las creencias herejes que sostuvo antes de su conversión religiosa. En las concepciones intrasubjetivas, al hablar de los diferentes «yoes» dentro de un «yo» (empírico y previamente individualizado) no lo hacemos sólo metafóricamente sino algunas veces con genuina relevancia moral y práctica.

Aunque Rawls no rechaza tales nociones de manera explícita, las niega implícitamente cuando asume que a cada ser humano individual le corresponde un sistema de intereses único, y que el utilitarismo fracasa como ética social al aplicar erróneamente a la sociedad los principios de elección adecuados para un solo hombre. Dado que da por sentado que cada individuo está conformado por un sistema de intereses, y sólo uno, el problema de fusionar tales intereses no surge en el caso individual, y la conducta de una persona hacia sí misma puede regirse adecuadamente por el principio de prudencia racional. «Una persona actúa de manera correcta, al menos cuando otros no resultan afectados, cuando trata de obtener el mayor beneficio posible y de promover sus fines racionales» (41). Mientras que la sociedad consiste en una pluralidad de sujetos, y por lo tanto requiere de la justicia, en la moral privada

parece bastar con el utilitarismo; cuando no hay otras personas involucradas, estoy libre para maximizar mi bien sin referencia al principio de lo justo.[8] Aquí nuevamente Rawls se aparta del pensamiento de Kant, quien pone énfasis en el concepto de «deber necesario ante uno mismo», y aplica la categoría del bien tanto a la moral privada como a la pública (Kant 1785: 37, 42, 45, 47, 50).

El supuesto de la posición original se opone por anticipado a toda concepción del bien que requiera una autocomprensión más o menos expansiva, y en particular a la posibilidad de una comunidad en el sentido constitutivo. En la perspectiva de Rawls, un sentido de comunidad describe un objetivo posible de un «yo» anticipadamente individualizado, y no un elemento o componente de su identidad como tal. Esto garantiza su rango subordinado. Dado que «la unidad esencial del "yo" es facilitada ya por la concepción de lo justo» (622), la comunidad debe encontrar su virtud como un competidor más entre otros, dentro del marco definido por la justicia, y no como una justificación rival del marco en sí. La cuestión se vuelve entonces determinar si los individuos que abrazan objetivos comunitarios pueden adoptarlos dentro de una sociedad bien ordenada, definida anticipadamente por los principios de la justicia, y no establecer si una sociedad bien ordenada es en *sí misma* una comunidad (en el sentido constitutivo). «Sin duda existe un objetivo colectivo con el respaldo del poder del Estado para la totalidad de la sociedad bien ordenada, una sociedad justa dentro de la cual el concepto compartido de la justicia está públicamente reconocido: pero dentro de este marco, pueden adoptarse objetivos comunitarios, y posiblemente una vasta mayoría de las personas lo hagan» (Rawls 1975: 550).

Podemos ver más claramente la relación entre la teoría de la persona y su postulado de la primacía de la justicia. Dado

8. En esta argumentación sobre la racionalidad deliberativa, Rawls reconoce poco menos que una dimensión intrasubjetiva y admite el concepto de lo justo como una restricción de la elección moral privada: «El que rechaza igualmente las exigencias de su «yo» futuro y los intereses de los otros no sólo es irresponsable con respecto a ellos, sino también en relación con su propia persona. No se ve a sí mismo como un ser que permanece. Ahora bien, considerado de este modo, el principio de responsabilidad ante sí mismo se asemeja a un principio del derecho. [...] La persona en un momento dado, por así decirlo, no debe poder lamentarse a causa de unas acciones que realice en otro momento» (467).

que los valores y fines de una persona son siempre atributos y nunca componentes del «yo», un sentido comunitario resulta siempre un atributo y nunca un componente de una sociedad bien ordenada. En tanto el «yo» es anterior a los objetivos que afirma, una sociedad bien ordenada, definida por la justicia, es anterior a los objetivos —comunitarios o de otro tipo— que profesen sus miembros. Es en este sentido, tanto moral como epistemológico, en el que la justicia es la primera virtud de las instituciones sociales.

Habiendo completado nuestra reconstrucción de la concepción de la persona de Rawls, queda por evaluar esta concepción y la ética deontológica que debe sustentar. Hemos visto que los supuestos que contiene la posición original son fuertes y de largo alcance en lugar de débiles e inocuos, aunque no por las razones que sugiere la crítica empirista. Estos supuestos no admiten todos los fines, sino que eliminan anticipadamente cualquier fin cuya adopción o consecución pueda comprometer o transformar la identidad del «yo», y en particular rechazan la posibilidad de que el bien de la comunidad consista en una dimensión constitutiva de este tipo.

Aunque no se puede decir que los principios de Rawls «se derivan de una doctrina que no impone restricciones anteriores» sobre las concepciones del bien, bien se puede alegar que las concepciones que se excluyen son en cierta forma dispensables, y que es posible justificar la justicia y llegar a una concepción de una sociedad ordenada sin ellas. La teoría de la justicia de Rawls es justamente un intento de este tipo. A fin de evaluarla, debemos descender de la metaética deontológica para considerar los principios de primer orden. En el siguiente capítulo, sostendré que la concepción de la persona de Rawls no puede sustentar ni su teoría de la justicia ni una justificación plausible de nuestras capacidades de agencia y de autorreflexión: la justicia no puede ser primaria en el sentido que lo requiere la deontología, ya que no podemos considerarnos coherentemente a nosotros mismos como la clase de seres que la ética deontológica nos exige que seamos.

II

La posesión, el mérito y la justicia distributiva

Después de haber aclarado el rango de los supuestos motores de Rawls, podemos proceder a un análisis comparativo de sus teorías de la persona y de la justicia para encontrar la conexión entre ambas. De esta forma podemos trabajar dentro del argumento que parte del equilibrio reflexivo, preguntándonos si la teoría de la persona que contiene la posición original se corresponde con los principios de justicia a los que debe dar forma y reflejar. Resulta de especial interés a este fin el principio de diferencia, que permite sólo aquéllas desigualdades que funcionan en beneficio de los miembros menos privilegiados de la sociedad. Encontraremos en este capítulo que una defensa adecuada del principio de diferencia presupone necesariamente una concepción de la persona que no está al alcance de los supuestos deontológicos, dado que no podemos ser sujetos para quienes es primaria la justicia y también ser sujetos para quienes el principio de diferencia es un principio de la justicia. Uno de los ejes centrales pasa por el papel del mérito en la justicia distributiva, y el concepto de la posesión que implica. Para explorar estos temas, comenzaremos comparando las perspectivas de Rawls con diversas teorías distributivas que compiten entre sí, y en particular la teoría contraria, aunque en algunos aspectos llamativamente semejante, sostenida por Robert Nozick (1974).

Del libertarianismo al igualitarismo

Desde un punto de vista político práctico, las posiciones de Rawls y de Nozick están claramente contrapuestas. Rawls, el liberal del estado de bienestar, y Nozick, el conservador libertario, definen entre ellos las más claras alternativas que ofrece la agenda política de los EE.UU., al menos con respecto a los temas que involucran la justicia distributiva. Y sin embargo, desde un punto de vista filosófico, tienen mucho en común. Ambos definen sus posiciones en oposición explícita al utilitarismo, que rechazan sobre la base de que éste niega la distinción entre las personas. Ambos ofrecen en su lugar una ética basada en derechos, que se supone garantiza la libertad de los individuos de forma más completa. Aunque la descripción de los derechos que hace Nozick debe mucho a Locke, ambos apelan al precepto kantiano de tratar a cada persona como un fin y no meramente como un medio, y buscan los principios de la justicia que los incorporen. Ambos niegan que exista entidad social alguna por encima o más allá de los individuos que la componen. Como escribe Nozick, haciendo eco de Rawls tanto en principio como en retórica:

> Las restricciones indirectas a la acción [es decir, las prohibiciones no calificadas] reflejan el principio kantiano subyacente de que los individuos son fines, no meramente medios [...] Las restricciones indirectas expresan la inviolabilidad de otras personas. Pero, ¿por qué no se puede infringir esta inviolabilidad por un bien social mayor? Cada uno de nosotros, individualmente, escogemos, algunas veces, padecer algún dolor o un sacrificio para obtener un beneficio mayor o para evitar un daño más grande. [...] Sin embargo, no hay ninguna *entidad social* con un bien, la cual soporte algún sacrificio por su propio beneficio. Hay sólo personas individuales, diferentes personas individuales, con sus propias vidas individuales. Usar a uno de estos individuos en beneficio de otros es usarlo a él y beneficiar a otros. Nada más. [...] Usar a una persona en esta forma no respeta, ni toma en cuenta suficientemente, el hecho de que es una persona separada, que ésta es la única vida que tiene. (1974: 43-45)
>
> Sostengo que las restricciones morales indirectas sobre lo que podemos hacer reflejan el hecho de nuestras existencias separadas. Estas restricciones reflejan el hecho de que ningún acto moralmente compensador puede tener lugar entre nosotros; no hay

nada que moralmente prepondere sobre una de nuestras vidas en forma que conduzca a un bien *social* general superior. No hay ningún sacrificio justificado de alguno de nosotros por los demás. (1974: 45)

Ambos teóricos enfatizan lo que Rawls denomina «la pluralidad y singularidad de las personas» y lo que Nozick llama «el hecho de nuestras existencias separadas». Éste es el hecho moral central que el utilitarismo niega y que la ética individualista, basada en los derechos, afirma. Acerca de este hecho moral, y de la importancia de los derechos, Rawls y Nozick están enfáticamente de acuerdo. Y sin embargo Rawls llega a una teoría de la justicia en la cual las desigualdades sociales y económicas son permisibles solamente en tanto beneficien a los menos privilegiados, mientras que Nozick mantiene que la justicia consiste únicamente en intercambios y transferencias voluntarios, excluyendo completamente las políticas redistributivas. ¿Cómo pueden llegar ambas teorías a una divergencia tan acentuada? Afortunadamente, el punto de divergencia puede ubicarse con cierta precisión, ya que Rawls, al desarrollar su segundo principio de la justicia (el que contiene el principio de diferencia), desarrolla una línea de razonamiento que parte de una posición similar a la de Nozick y finaliza con la propia.

Rawls considera tres principios posibles por los cuales la distribución de beneficios económicos y sociales puede regularse o determinarse: la libertad natural (similar a la «teoría de los derechos» de Nozick), la igualdad liberal (cercana a una meritocracia estándar) y la igualdad democrática (basada en el principio de diferencia). El sistema de la libertad natural define como justo cualquier resultado distributivo que surge de una economía de mercado eficiente en la cual prevalece una igualdad de oportunidades formal (v.g., legal), de forma tal que el acceso a los puestos de trabajo esté abierto a quienes poseen las cualidades relevantes. Rawls considera que este principio es inadecuado sobre la base de que la distribución que autoriza tiende simplemente a reproducir la distribución inicial de talentos y ventajas competitivas; aquéllos substancialmente más calificados terminarán con las porciones más sustanciales, y aquéllos con magras aptitudes terminarán con magros resultados. Cuando el resultado tiende simplemente a reproducir la distribución inicial, sólo es posible

designarlo como justo si se hace la presunción adicional de que los atributos iniciales están distribuidos con justicia. Pero no es posible afirmar esta presunción; «la distribución inicial del activo para cualquier período está fuertemente influenciada por contingencias naturales y sociales», y como tal, no es ni justa ni injusta sino simplemente arbitraria. Y ya que nada garantiza la justicia de la distribución inicial de atributos, el instalarlos en el nombre de la justicia es incorporar la arbitrariedad de la suerte, nada más. «Intuitivamente, la injusticia más obvia del sistema de libertad natural es que permite que las porciones distribuidas se vean incorrectamente influenciadas por estos factores que desde el punto de vista moral son tan arbitrarios» (94).

El principio de igualdad liberal busca solucionar la injusticia de la libertad natural yendo más allá de la igualdad de oportunidades formal y corrigiendo, de ser posible, las desventajas sociales y culturales. El objetivo es una suerte de «meritocracia justa» en la cual las desigualdades sociales y culturales son mitigadas por la igualdad de oportunidades en la educación, ciertas políticas redistributivas, y otras reformas sociales. En el principio de la igualdad liberal, el ideal es brindar un «punto de partida» igualitario para todos, a fin de que aquéllos con talento y capacidades innatas similares y una voluntad similar de ejercitarlos tengan «las mismas perspectivas de éxito, independientemente de su posición inicial en el sistema social. En todos los sectores de la sociedad debería haber, en términos generales, las mismas perspectivas de cultura y de éxito para todos los que se encuentran igualmente motivados y dotados. Las expectativas de aquellos con las mismas capacidades y aspiraciones no deberían verse afectadas por sus clases sociales» (94-95).

Pero aunque la igualdad liberal representa una evolución en comparación con el sistema de la libertad natural, «intuitivamente sigue pareciendo insuficiente». La igualdad de oportunidades, no importa cuán completa sea, es un embate demasiado débil contra la arbitrariedad de la suerte.

> [...] Aun si funcionase a la perfección eliminando la influencia de las contingencias sociales, de todas maneras permitiría que la distribución de la riqueza y del ingreso fuesen determinadas por la distribución natural de capacidades y talentos. Dentro de los límites permitidos por las condiciones subyacentes, las porciones distribui-

das se deciden conforme al resultado de una lotería natural; y desde una perspectiva moral este resultado es arbitrario. No hay mejor razón para permitir que la distribución del ingreso y la riqueza sea resuelta en función de las capacidades naturales, que para que lo sea a favor de las contingencias sociales e históricas. (95-96)

Una vez tomamos conciencia de la arbitrariedad del reparto inicial de atributos que determina nuestras perspectivas de vida en general, estamos destinados por reflejo a vernos tan afectados por la influencia de las contingencias naturales como por las sociales y culturales. «Desde un punto de vista moral, ambas parecen igualmente arbitrarias» (97). El mismo razonamiento que nos lleva a favorecer una «meritocracia justa» (como la de la igualdad liberal) sobre una igualdad de oportunidades puramente formal (como en la libertad natural), naturalmente nos lleva adelante en la búsqueda de lo que Rawls denomina una concepción democrática. Pero parece claro que la concepción democrática no puede encontrarse en una simple extensión del principio de igualdad de oportunidades. Por una parte, sería virtualmente imposible extender las oportunidades de manera tan completa como para erradicar incluso aquéllas desigualdades que se derivan únicamente de las condiciones sociales y culturales. La institución de la familia, por sí sola hace que «en la práctica (sea) imposible asegurar a los igualmente dotados iguales probabilidades culturales y de superación [...]» (96). Pero incluso si la educación compensatoria y otras reformas pudieran corregir completa o al menos parcialmente las carencias sociales y culturales, resulta difícil, si no vagamente inadmisible, imaginar qué tipo de políticas sociales serían necesarias para «corregir» de manera similar los avatares naturales de la fortuna. Lo que se necesita, por lo tanto, es una concepción que anule los *efectos* de estas diferencias y que al mismo tiempo reconozca la imposibilidad de erradicarlas.

Algunos críticos, y particularmente aquéllos que se oponen al principio de la igualdad democrática, describen el siguiente paso lógico como un desplazamiento de la igualdad de oportunidades a la igualdad de resultados. En su opinión, cualquier teoría de la justicia que rechace el concepto de la meritocracia por la arbitrariedad moral de sus consecuencias distributivas debe necesariamente estar comprometida con una forma de igualdad niveladora que exigiría ajustes constantes de las por-

ciones distributivas para corregir las diferencias persistentes de talento y habilidad innatos (Bell 1973: 441-3). Pero la igualdad de resultados no es de ninguna manera la única alternativa democrática al sistema de la meritocracia, ni es el principio adoptado por Rawls. El principio de diferencia no es sinónimo de la igualdad de resultados, ni requiere la nivelación de todas las diferencias entre las personas. «No hay razón para ignorar y mucho menos para eliminar estas distinciones», escribe Rawls, «hay otra manera de tratarlas» (124). La manera que sugiere Rawls no es erradicar la atribución desigual de aptitudes sino organizar el esquema de cargas y beneficios para que los menos privilegiados puedan compartir los recursos de los afortunados. Éste es el arreglo al que apunta el principio de diferencia. Define como justas únicamente aquellas desigualdades sociales y económicas que funcionan para el beneficio de los miembros menos privilegiados de la sociedad. Tomado junto con el principio de la idoneidad para el acceso a los puestos de trabajo privados y públicos, el principio de diferencia define la concepción de Rawls de la igualdad democrática.

El principio de diferencia no es simplemente una versión más completa del principio de la igualdad de oportunidades; ataca el problema de la arbitrariedad de una manera fundamentalmente distinta. En lugar de transformar las condiciones bajo las cuales pongo en práctica mis talentos, el principio de diferencia transforma la base moral en la que me baso para exigir los beneficios que se derivan de ellos. Ya no se me considera el único propietario de mis aptitudes, o el receptor privilegiado de las ventajas que producen. «El principio de diferencia representa, en efecto, un acuerdo en el sentido de considerar la distribución de talentos naturales, en ciertos aspectos, como un acervo común, y de participar en los mayores beneficios económicos y sociales que hacen posibles los beneficios de esa distribución» (124). De esta manera, el principio de diferencia reconoce la arbitrariedad de la suerte al afirmar que yo no soy realmente el propietario sino simplemente el custodio o guardián de los talentos y capacidades que residen en mí, y como tal no tengo ningún derecho especial sobre los frutos de su ejercicio.

> Aquellos que han sido favorecidos por la naturaleza, quienes quiera que fuesen, pueden obtener un provecho de su buena suerte sólo en la medida en que mejoren la situación de los no favore-

cidos. Los favorecidos por la naturaleza no podrán obtener ganancias por el mero hecho de estar más dotados, sino solamente para cubrir los costos de su entrenamiento y educación y para usar sus dones de manera que también ayuden a los menos afortunados. Nadie merece una mayor capacidad natural ni tampoco un lugar inicial más favorecido en la sociedad. (124)

Al considerar la distribución de talentos y atributos como un acervo común en lugar de una posesión individual, Rawls obvia la necesidad de «nivelar» las aptitudes a fin de solucionar la arbitrariedad de las contingencias sociales y naturales. Cuando «los hombres convienen en aprovecharse de los accidentes de la naturaleza y de las circunstancias sociales» (125), resulta menos importante el hecho de que sus suertes, individualmente, puedan variar. Por esto aunque el principio de diferencia tiende a «revertir el sesgo de las contingencias en dirección a la igualdad», «no exige que la sociedad trate de nivelar las desventajas como si se esperara que todos fueran a competir sobre una base equitativa en la misma carrera» (123).

Rawls reconoce que el principio de diferencia y en particular la noción de las aptitudes como un «acervo común» rompe con la concepción tradicional del mérito individual. «Podría objetarse que aquellos que están mejor situados merecen las mayores ventajas que pudieran adquirir para sí [...] ya sea que obtengan estas ventajas en formas que beneficien a otros o no» (126). La respuesta que da Rawls es que esta concepción del mérito individual es un error, como sugiere el argumento general de la arbitrariedad. «No merecemos el lugar que tenemos en la distribución de dones naturales, como tampoco nuestra posición inicial en la sociedad» (126). La pretensión de que una persona merece al menos lo que ella logra a través de su propio esfuerzo es más plausible en lo intuitivo, pero incluso la voluntad de esforzarse concienzudamente puede estar determinada en gran medida por circunstancias sociales y naturales. «Igualmente problemático es el que merezcamos el carácter superior que nos permite hacer el esfuerzo por cultivar nuestras capacidades, ya que tal carácter depende, en buena medida, de condiciones sociales conocidas y afortunadas en la niñez, por las cuales no puede pretenderse crédito alguno. La noción de mérito no puede aplicarse aquí» (126).

Esto no constituye una negación del papel que juegan los derechos individuales. Ciertos tipos de derechos son coheren-

tes con el principio de diferencia, pero es necesario distinguir entre el mérito y las expectativas legítimas. Dado que sirve al interés general que yo cultive y ejercite (algunos de) los talentos y dones que poseo, en lugar de dejar que permanezcan latentes, la sociedad está estructurada típicamente para brindar recursos para su cultivo e incentivos para su ejercicio. Evidentemente, tengo derecho a mi porción de tales beneficios cuando me he hecho acreedor a ellos bajo los términos especificados. Lo que resulta importante destacar, sin embargo, es que *estas* demandas cumplen con las expectativas legítimas creadas por las instituciones, que están diseñadas para estimular mis esfuerzos, y no con un derecho o demanda de mérito en virtud de las cualidades que poseo.

> Ahora bien, es verdad que, dado un sistema justo de cooperación como marco de las reglas públicas y las expectativas creadas por él, aquellos que, con el propósito de mejorar su condición, hayan hecho lo que el sistema anuncia que recompensará, están facultados a ver satisfechas sus expectativas. En este sentido los más afortunados tienen derecho a su mejor situación; sus demandas son expectativas legítimas establecidas por las instituciones sociales y la comunidad está obligada a satisfacerlas. Pero este sentido de lo merecido es el de lo otorgado. Presupone la existencia de un esquema cooperativo que funciona, y nada tiene que ver con el problema de si el esquema mismo habrá de ser diseñado de acuerdo con el principio de diferencia o con algún otro criterio. (126)

Aunque esté *facultado* para los beneficios que responden a mis expectativas legítimas, no los *merezco*, por dos razones. En primer lugar, dado el supuesto del acervo común, yo no *poseo* en realidad los atributos que son causa de los beneficios, o si los poseo, es solamente en un sentido débil y accidental y no en el sentido fuerte constitutivo, y este sentido de la posesión no es el adecuado para establecer el mérito en el sentido fuerte, preinstitucional. En segundo lugar, mientras que estas reglas me facultan a recibir mi porción justa, no estoy facultado a exigir que las reglas en vigencia sean *éstas*, las que recompensan *estos* atributos, y no otras. Por estas razones, «es incorrecto que los individuos con mayores dones naturales y con el carácter superior que ha hecho posible su desarrollo tengan derecho a un esquema cooperativo que les permita obtener aún más beneficios en formas que no contribuyan al beneficio de los demás» (126).

La meritocracia y el principio de diferencia

Antes de avanzar hacia la consideración de la respuesta de Nozick a la pregunta que plantea Rawls y hacia su defensa de la libertad natural, puede resultar útil aclarar algunos de los contrastes entre el principio de diferencia (en la «igualdad democrática» de Rawls) y la concepción de la meritocracia (en la «igualdad liberal»). Quizás la diferencia más notable reside en el papel que juega el mérito individual, que es central en la concepción meritocrática y se encuentra ausente, o al menos seriamente debilitado, en la justicia como imparcialidad.[9] En una «meritocracia justa», es decir, aquélla en la cual se han superado la discriminación y el sesgo de clase, quienes llegan a posiciones favorecidas se han ganado tal condición, y por lo tanto merecen las recompensas que ésta implica. Se asignan porciones distributivas desiguales en reconocimiento de logros superiores y no meramente para satisfacer expectativas legítimas. Como explica un defensor de la ética de la meritocracia:

> Una meritocracia está compuesta por aquellos que se han ganado su autoridad. [...] La meritocracia, en el contexto de mi utilización del término, es un énfasis en los logros individuales y el nivel logrado que confirman nuestros pares. [...] Mientras que todos los hombres tienen derecho a ser respetados, no todos tienen derecho a ser elogiados. La meritocracia, en el mejor sentido del término, está compuesta por aquellos que son merecedores de elogio. (Bell 1973: 453-4)

Una segunda diferencia, relacionada, se refiere a la distinción entre las ventajas genéticas y las culturales. Para Rawls, esta distinción es virtualmente irrelevante para las cuestiones de la justicia. En el supuesto meritocrático, por el contrario, la distinción es crucial: de ahí el intenso debate entre los que sostienen los ideales de la meritocracia sobre el papel de los factores genéticos y culturales en la determinación de la inteligencia y de las perspectivas de vida en general. Cuando la estructura de la justicia distributiva es considerada como

9. No incluyo aquellas versiones de la meritocracia que asignarían porciones distributivas con el propósito de crear incentivos y atraer únicamente a los talentos sobresalientes, sin referencia al valor moral de los receptores.

dependiente de la «igualdad de oportunidades» de todos los que compiten igualmente por recompensas en última instancia desiguales, la distinción entre los obstáculos genéticos y los sociales o culturales se vuelve central para cualquier evaluación del esquema. Cuanto más estrechamente se relaciona el éxito con factores hereditarios, puede presumirse que las instituciones sociales serán capaces o responsables de mejorar menos desigualdades, y menor parece ser el rango disponible para la clase de esfuerzo individual del que se presume depende el mérito.

> En la naturaleza de la meritocracia, como se ha concebido tradicionalmente, lo que resulta central para la evaluación de una persona es la relación supuesta entre el logro y la inteligencia, y entre la inteligencia y su medición por medio de la escala del *coeficiente intelectual*. La primera cuestión, por lo tanto, es qué determina la inteligencia.
>
> Todo esto hace que la cuestión de la relación entre inteligencia y herencia genética sea muy delicada. ¿La inteligencia se hereda en gran medida? ¿Se puede elevar la inteligencia nutriéndola? ¿Cómo se pueden separar la capacidad y el empuje innatos del perfeccionamiento logrado a través de la educación? (Bell 1973: 411)

Para la justicia como imparcialidad, el debate sobre los determinantes de la inteligencia y la gran literatura científica social que se ha producido resulta más o menos irrelevante. Una vez que se rechazan las nociones del mérito individual y la «igualdad de oportunidades» como las bases primordiales de las porciones distributivas, la distinción entre obstáculos genéticos y culturales para el éxito pierde gran parte de su interés moral. Una vez que se acuerda que la distribución de talentos es un acervo común, poco importa cómo han llegado algunos a residir en ti y otros en mí.

Los defensores de las concepciones meritocráticas no siempre son explícitos acerca de las *bases* de su distinción entre ventajas naturales y sociales, pero uno puede imaginar al menos dos argumentos posibles, uno moral y el otro de índole práctica. El primero afirmaría que las retribuciones genéticas son inviolables de una manera en que las características sociales o culturales no lo son, y que los atributos naturales de una persona son en cierta manera más esencialmente *suyos*,

más profundamente constitutivos de su identidad que sus atributos condicionados socialmente. Las diferencias innatas, por más arbitrarias que resulten, no son dispensables de la misma manera; son ellas, y no los resultados del condicionamiento social y cultural, las que marcan los rasgos en ausencia de los cuales yo no sería la persona particular que soy. Dentro de esta perspectiva, la cuestión no es el hecho de que yo merezca la inteligencia con la que nací, por ejemplo; lo que importa es que mi inteligencia innata es un hecho irreductible sobre *mí*, y la integridad de mi persona exige que no se altere, no importa cuán valioso sea el propósito social general.

Pero este argumento aparece como erróneo una vez que establecemos que aquellas cualidades que se consideran más plausiblemente como esenciales para la identidad de la persona (nuestro carácter, valores, convicciones básicas y lealtades más profundas, por ejemplo) a menudo son fuertemente influenciados por factores sociales y culturales, mientras que muchos factores naturales (tales como el color del pelo y otras características físicas triviales) son más evidentemente dispensables. Incluso aunque resulta válida cierta distinción entre las características esenciales y las meramente accidentales, no parece existir ninguna razón evidente por la que deba corresponderse con la distinción entre activos sociales y naturales. Por supuesto, la objeción de Rawls al argumento sería aún más fuerte, ya que su teoría de la persona implica que ninguna característica, ya sea social o natural, puede resultar esencial en este sentido. Incluso aquellos atributos, tales como el carácter y los valores de una persona, que intuitivamente parecen más cercanos a la definición de un «yo» esencial son descriptos como contingentes. Dado que el carácter de una persona «depende, en buena medida, de condiciones familiares y sociales afortunadas en la niñez, por las cuales no puede pretenderse crédito alguno» (126), nuestros valores son también accidentales. «El que tengamos una concepción del bien en lugar de otra no es relevante desde un punto de vista moral. Al adquirirla estamos influenciados por la misma clase de contingencias que nos lleva a eliminar el conocimiento de nuestro sexo y clase» (Rawls 1975: 537).

El argumento práctico distinguiría las desigualdades naturales de las sociales sobre la base de que las primeras son contumaces de una manera en que las segundas no lo son, y de que la sociedad puede por lo tanto considerarse la responsable

de compensar las desigualdades sociales pero no las naturales. En la medida en que la desigualdad resulta surgir de factores genéticos y no culturales, es menos lo que la sociedad «puede hacer al respecto». Dado un sistema justo, algunos avanzarán con mayor éxito que otros, y llega un punto en que incluso la sociedad más iluminada no puede hacer nada para alterar este hecho. En determinada instancia, incluso los reformistas más empecinados deben reconocer que la vida es injusta de una manera que ninguna institución social puede tener la esperanza de corregir. Las personas son diferentes, y esas diferencias, tarde o temprano, inevitablemente surgen a la luz, incluso (o quizás más certeramente) en una sociedad donde prevalece la igualdad de oportunidades. «Lo que es importante es que la sociedad, en la medida más completa posible, sea genuinamente abierta» (Bell 1973: 454).

A esto Rawls respondería probablemente que el papel de la sociedad está naturalmente limitado de esta forma sólo si asumimos que su única función en la promoción de la justicia consiste en sus esfuerzos por nivelar las desventajas de los menos afortunados para que puedan competir en forma más justa. Pero este supuesto no tiene en cuenta la elección social igualmente significativa implícita en los objetivos que siguen las instituciones y los atributos que recompensan en el proceso. Incluso si la vasta mayoría de las diferencias entre las personas resultara ser genética en lugar de cultural, todavía la sociedad tendría que determinar *cuáles* de esas diferencias (si es que alguna) deberían constituirse adecuadamente como la base de las porciones distributivas diferenciales. Evidentemente, si el objetivo de las instituciones sociales se toma como fijo, como por ejemplo la maximización del producto social total, los defensores de una «meritocracia justa» tienen razón, en el sentido de que la única cuestión pendiente para la justicia es lo bien equipadas que estén las personas por la sociedad para contribuir a este propósito y para cosechar los beneficios de su contribución. Pero inmediatamente surge la cuestión de por qué *ese* fin determinado debe ser primario, incluso aunque reconozcamos que sólo puede prevalecer a expensas de la injusticia social. En resumen, una sociedad no sólo equipa o no equipa a sus miembros con las retribuciones relevantes a su propósito colectivo, sino que también define a través de su estructura institucional la naturaleza de tal propósito: de esto se deduce que también determina los atributos favorecidos e

instalados como la base de las porciones distributivas. En una pasaje de cierta elocuencia, Rawls afirma:

> A la luz de estas observaciones podemos rechazar la afirmación de que la ordenación de las instituciones sea siempre defectuosa porque la distribución de los talentos naturales y las contingencias de la circunstancia social son injustas, y que esta injusticia se transmite inevitablemente a los acuerdos humanos. Esta reflexión es presentada en ocasiones como excusa para tolerar la injusticia, como si el negarse a aceptar la injusticia fuera comparable con la incapacidad para aceptar la muerte. La distribución natural no es ni justa ni injusta, como tampoco es injusto que las personas nazcan en una determinada posición social. Éstos son hechos meramente naturales. Lo que puede ser justo o injusto es el modo en que las instituciones actúan respecto a estos hechos. (124-125)

Estas consideraciones nos llevan naturalmente a un tercer contraste entre la concepción meritocrática y la democrática, con respecto a la relación entre el valor de varios activos y atributos por una parte, y las instituciones que los aprecian y recompensan por la otra. En la concepción meritocrática, las instituciones sociales tienden a recompensar ciertos atributos y no otros. Las cualidades que exige un grupo de instituciones tienen un valor antecedente a su valuación institucional, y por lo tanto brindan una prueba independiente de la justicia de las instituciones mismas. Los acuerdos institucionales que recompensan cualidades nobles en lugar de cualidades bajas son, debido a *esto*, los más valiosos, aparte de otras consideraciones relevantes a la justicia, tales como el propósito que promueven.

En la perspectiva de Rawls, las instituciones no se consideran de esta manera, ya que las virtudes que las restringirían esperan a su vez una definición institucional. «El concepto de valor moral no proporciona un primer principio de justicia distributiva» (351). Ya que ninguna virtud tiene un rango antecedente o preinstitucional, el diseño de las instituciones es abierto con respecto a las cualidades que pueden recompensar. Como resultado, el valor intrínseco de los atributos que una sociedad busca y recompensa no puede establecerse como una medida para evaluar su justicia, ya que, para empezar, su valor sólo surge a la luz de los acuerdos institucionales. El rechazo de las nociones preinstitucionales de la virtud que for-

mula Rawls refleja la prioridad de lo justo sobre lo bueno y su negativa a elegir por anticipado entre concepciones del bien que compiten entre sí. Una sociedad cazadora que recompensa la velocidad para la carrera por encima la elocuencia (que sería apreciada en una sociedad litigiosa, por ejemplo), no es menos justa o valiosa debido a *ello,* ya que no hay bases antecedentes sobre las cuales pueda demostrarse que la rapidez en la carrera es más o menos virtuosa que la rapidez en la réplica. La prioridad de las sociedades justas con respecto a la virtud y al valor moral brinda una segunda razón por la que no puedo decir que merezco los beneficios que se derivan de mis atributos naturales. Para que yo merezca los beneficios asociados con «mi» inteligencia superior (por ejemplo) es necesario tanto que yo posea mi inteligencia (en un sentido no arbitrario de la posesión) como que tenga un derecho (en el sentido fuerte, preinstitucional del derecho) a que la sociedad aprecie la inteligencia por encima de otros bienes. Pero en la concepción de Rawls, no se da ninguna de estas condiciones. El argumento que conduce desde la arbitrariedad al acervo común se opone a la primera, y la precedencia de las instituciones sobre el valor moral niega la segunda.

La defensa del acervo común

En el desarrollo de nuestra crítica a la teoría de la persona y al principio de diferencia de Rawls, la una a la luz del otro, debemos tomar como puntos de partida dos líneas de la crítica que formula Nozick a la justicia como imparcialidad. La primera ataca el principio de diferencia y en particular la noción de la posesión sobre la que se basa, y la segunda defiende una versión de la libertad natural al fijarse sobre las nociones de mérito y derecho. Comparando de esta manera a ambos autores podemos evaluar la concepción de Rawls y también poner en claro algunas de las similitudes y diferencias entre sus respectivas teorías distributivas.

La objeción central de Nozick al argumento que produce el principio de diferencia se centra en la perspectiva de Rawls de que la distribución de talentos naturales debe considerarse como una posesión «común» o «colectiva» que debe compar-

tirse dentro de la sociedad como un todo. En palabras de Rawls:

> El principio de diferencia representa, en efecto, un acuerdo en el sentido de considerar la distribución de talentos naturales, en ciertos aspectos, como un acervo común, y de participar en los mayores beneficios económicos y sociales que hacen posibles los beneficios de esa distribución. (124)
>
> Como ya he señalado, los dos principios son equivalentes al considerar la distribución de capacidades naturales, en algunos aspectos, como una ventaja colectiva, haciendo que los más afortunados sólo se beneficien ayudando a los menos favorecidos. (208)

Rawls cree que la noción del acervo común tal como está incorporada en el principio de diferencia expresa el ideal del respeto mutuo que el liberalismo deontológico pretende afirmar.

> [...] al disponer las desigualdades para beneficio recíproco y al abstenerse de la explotación de las contingencias de la naturaleza y de las circunstancias sociales dentro de un marco de libertades iguales, las personas expresan su mutuo respeto en la constitución misma de la sociedad. [...] Otra manera de expresar esto es decir que los principios de la justicia reflejan en la estructura básica de la sociedad el deseo que tienen los hombres de no tratarse como medios sino únicamente como fines en sí mismos. (208-209)

Nozick, por el contrario, alega que considerar las ventajas naturales de las personas como un bien común es precisamente contradecir todo lo que afirma el liberalismo deontológico al enfatizar la inviolabilidad del individuo y la diferencia entre las personas.

> Las personas diferirán según cómo consideren los talentos naturales como dotes comunes. Algunos se quejarán haciendo eco a Rawls, en contra del utilitarismo, de que esto «no toma seriamente la distinción entre personas»; y se preguntarán si cualquier reconstrucción de Kant que trata las capacidades y talentos de las personas como recurso para otros puede ser adecuada. «Los dos principios de la justicia [...] excluyen hasta la tendencia de considerar a los hombres como medios para el bienestar de otros». Sólo si se insiste *mucho* en la distinción entre los hombres y sus talentos, dotes, habilidades y particularidades especiales. (Nozick 1974: 228)

Nozick llega al núcleo de la teoría del sujeto de Rawls. Como hemos visto, Rawls pone un énfasis realmente fuerte sobre la distinción entre el «yo» y sus distintas posesiones. La severidad de esta distinción, a pesar de lo problemática que pueda resultar, está cuidadosamente adaptada a las exigencias del proyecto deontológico como un todo; acepta la prioridad del «yo» sobre sus fines, que a su vez sustenta la prioridad de lo justo y la primacía de la justicia. Otro rasgo de esta concepción es que permite la defensa siguiente contra la objeción de Nozick al principio de diferencia.

El considerar a la distribución de talentos naturales como un acervo común no viola la diferencia entre las personas ni considera a algunas como medios del bienestar de las otras, porque no son las *personas* sino sus *atributos* los que se usan como medios para el bienestar de los demás. Decir que de alguna manera soy yo el que sufre una violación o un abuso cuando «mi» inteligencia o incluso mi esfuerzo se usan para el beneficio común es confundir al «yo» con sus atributos, que son contingentes y para nada esenciales (no esenciales para que yo sea el ser particular que soy). Sólo una teoría de la persona que mantenga que estos bienes son *elementos constitutivos* esenciales en lugar de *atributos* alienables del «yo» podría considerar que el compartir los atributos equivale a usar al *sujeto* como medio para los fines de otro. Pero en la concepción de Rawls todos los bienes son contingentes y en principio separables del «yo», cuya prioridad se asegura a través de su habilidad para retroceder constantemente frente al remolino de las circunstancias. Ésta es la característica que preserva su identidad, al asegurar su invulnerabilidad frente a la transformación por la experiencia.

Aunque esta defensa elude la inconsistencia, invita enseguida a una objeción de incoherencia relacionada, ya que si Rawls debe invocar la distinción entre el «yo» y sus posesiones en esta forma profundizada, surge inmediatamente la cuestión de si, al intentar evitar un sujeto radicalmente situado, no retrocede al extremo opuesto proponiendo un sujeto radicalmente incorpóreo. Nozick formula la objeción: «Si alguna concepción coherente de persona queda cuando se insiste de esta manera en la distinción, esto es una pregunta abierta. El por qué nosotros, llenos (como estamos) de rasgos particulares, debemos estar contentos de que (sólo) los hombres así purificados dentro de nosotros no sean considerados como medios, tampoco es claro» (Nozick 1974:224).

Así anticipa la defensa de Rawls y muestra que es demasiado artificiosa para redimir la teoría. La noción de que sólo *mis atributos* son utilizados como medios, y no *yo*, amenaza con vulnerar lo plausible e incluso lo coherente de la distinción misma que invoca. Sugiere que en el principio de diferencia, podemos tomar seriamente la distinción entre personas solamente tomando metafísicamente la distinción entre una persona y sus atributos. Pero esto tiene como consecuencia que nos deja con un sujeto tan privado de características empíricamente identificables (tan «purificado», en las palabras de Nozick), que nos recuerda al sujeto trascendente o incorpóreo de Kant que Rawls justamente se propone evitar. Parecería que Rawls puede defenderse de la acusación de inconsistencia solamente al precio de la incoherencia, y que la objeción de Nozick al principio de diferencia es por lo tanto exitosa.

Pero Rawls tiene a su disposición una defensa alternativa, que no fue anticipada por Nozick. Aunque rescata el principio de diferencia de su fundamentación en una concepción aparentemente incorpórea del sujeto, lo hace a expensas de otros aspectos de la doctrina de Rawls, por lo que es probable que el mismo Rawls se resistiera a ello. Sin embargo, trataré de demostrar que el principio de diferencia solamente puede sostenerse adoptándolo de alguna manera. En esta segunda defensa, Rawls podría negar que el principio de diferencia me utilice a mí como un medio para los fines de otros, no porque son mis *activos* y no mi *persona* los que están siendo utilizados, sino cuestionando el sentido en el cual aquellos que se benefician de «mis» atributos pueden describirse adecuadamente como «otros». Mientras que la primera defensa enfatiza la distinción entre el «yo» y sus atributos, la segunda califica la distinción entre el «yo» y los otros al proponer que, en ciertas circunstancias morales, la descripción relevante de un «yo» puede incluir más de un ser humano empíricamente individualizado. La segunda defensa vincula la noción de acervo común con la posibilidad de un sujeto común de la posesión. Apela, en resumen, a una concepción intersubjetiva del «yo».

La única salida aparente a las dificultades que plantea Nozick consiste en postular que el principio de diferencia involucra a Rawls en una concepción intersubjetiva que por otro lado rechaza. También sirve para iluminar un supuesto no debatido de la teoría del sujeto de Rawls. Como hemos visto, Rawls concibe al «yo» como un sujeto de la posesión, limitado antici-

padamente, y dado con prioridad a sus fines, y da por supuesto, además, que los límites del sujeto se corresponden sin problemas con los límites corporales entre los seres humanos individuales. Pero este postulado nunca es defendido, sino sólo supuesto, y su carácter cuestionable se esconde probablemente tras la afinidad que tiene con nuestra perspectiva de sentido común reflexivo sobre el tema.[10] El énfasis que Rawls pone en la pluralidad como rasgo esencial de la sociedad humana puede pensarse como un respaldo de este supuesto, pero sólo puede establecer que *algún* principio de pluralidad o diferenciación es esencial para una explicación del sujeto humano, que no necesariamente es de carácter físicamente corporal; ni tampoco demuestra que el número de esta pluralidad deba corresponderse en todos los casos con el número de seres humanos individuales que hay en el mundo.

En todo caso éste es el supuesto que debe ceder para poder superar la objeción de Nozick al acervo común. Si el principio de diferencia tiene que evitar utilizar algo como medio para el fin de otros, sólo puede hacerlo bajo circunstancias en las que el sujeto de la posesión sea un «nosotros» en lugar de un «yo», circunstancias que implican a su vez la existencia de una comunidad en el sentido constitutivo.

La conclusión de que la teoría de Rawls se apoya implícitamente en una concepción intersubjetiva que oficialmente él rechaza encuentra sustento también en la argumentación del mérito y la justificación que la sigue, y se insinúa en varias huellas de lenguaje intersubjetivo que podemos encontrar en todo el texto. Este lenguaje aparece por primera vez en la argumentación del principio de diferencia cuando, como hemos visto, la distribución de talentos naturales se describe alternativamente como un acervo «común», «colectivo» o «social» que debe utilizarse «para beneficio común» (124, 130, 208). «En la justicia como imparcialidad las personas acuerdan compartir el destino común» (125). Resuelven no tomar seriamente las diferencias entre las personas como la base para perspectivas de vida diferenciadas, sobre la base de que tales diferencias surgen de factores arbitrarios desde el punto de vista moral.

10. Aun así, de la crítica de Rawls al utilitarismo se desprende como particularmente sorprendente el hecho de que no defienda su principio de individualización en forma más completa, ya que es precisamente en este tema donde afirma que el utilitarismo se equivoca al no reconocer o, por lo menos, no tomar en serio la distinción entre las personas.

En su argumentación sobre la idea de la unión social, Rawls lleva su lenguaje intersubjetivo desde el acervo común hacia los fines y propósitos comunes, y habla de los seres humanos que realizan su naturaleza común en una retórica que se acerca peligrosamente a la teleológica. En su explicación de la unión social, Rawls se aparta de los temas distributivos para ocuparse de la autorrealización, e intenta demostrar que la justicia como imparcialidad puede brindar una interpretación de la sociabilidad humana que no es ni trivial ni puramente instrumental. «Así, los seres humanos tienen, de hecho, objetivos finales compartidos, y valoran sus instituciones y actividades comunes como buenas en sí mismas» (577). Las características de una unión social son «objetivos finales compartidos y en actividades comunes valoradas por sí mismas» (580). Siguiendo a Humboldt, un liberal del siglo XIX de la tradición idealista alemana, Rawls afirma que «es a través de la unión social fundada en las necesidades y posibilidades de sus miembros como cada persona puede participar en la suma total de los valores naturales realizados de los otros. [...] *Sólo en una unión social se completa el individuo*» (578, 580n). Las uniones sociales se encuentran en una variedad de formas y tamaños; «van desde las familias y las amistades hasta asociaciones mucho más amplias. También carecen de límites de tiempo y de espacio, porque las que se encuentran muy separadas por la historia y por las circunstancias pueden, sin embargo, cooperar en la realización de *su naturaleza común*» [Los subrayados son míos] (582).

A través de sus dimensiones intersubjetivas, el principio de diferencia y la idea de la unión social se contraponen a los supuestos individualistas de dos maneras diferentes: el principio de diferencia al anular la arbitrariedad que surge cuando los atributos naturales se consideran una posesión individual, y la idea de la unión social al superar la parcialidad de las personas que aparece cuando se considera que los individuos están completos en sí mismos. En una unión social «los miembros de una comunidad participan de sus naturalezas recíprocas [...] [y] *el "yo" se realiza en las actividades de muchos* [yoes]». [Los subrayados son míos] (624).

Es una característica de la sociabilidad humana que no somos más que partes de lo que podríamos ser. Debemos cuidar de que los otros alcancen las excelencias que nosotros tenemos que dejar

a un lado, o de las que carecemos totalmente. La actividad colectiva de la sociedad, las numerosas asociaciones y la vida pública de la comunidad más amplia que las regula, mantienen nuestros esfuerzos y suscitan nuestra colaboración. Pero el bien alcanzado gracias a nuestra cultura común sobrepasa considerablemente nuestro trabajo, en el sentido de que *dejamos de ser simples fragmentos*: la *parte de nosotros mismos* que realizamos directamente se une a un ordenamiento más amplio y más justo, cuyos objetivos nosotros afirmamos. [Los subrayados son míos.] (583)

El fundamento de la noción del mérito

La noción de posesión conduce naturalmente a demandas de mérito y de derechos. El debate sobre qué es lo que poseen las personas, y en qué términos, tiene un impacto directo en la cuestión de qué es lo que merecen las personas o a qué tienen derecho en términos de justicia. Nos volvemos hacia estos temas de mérito y de derecho para considerar la segunda línea de la crítica de Nozick a la justicia como imparcialidad. Rawls rechaza los principios de la libertad natural y de la igualdad liberal sobre la base de que recompensan atributos y activos, los cuales, al ser arbitrarios desde un punto de vista moral, no podemos realmente decir que las personas merezcan, y adopta el principio de diferencia sobre la base de que anula esta arbitrariedad. Nozick ataca su línea de razonamiento alegando en primer término que la arbitrariedad no afecta al mérito, y en segundo, que aunque lo hiciera emergería como el resultado preferible de una versión de la libertad natural y no del principio de diferencia.

En términos de la posesión, la objeción de Rawls a la libertad natural y a la igualdad liberal señala que, bajo estos principios, se permite que las personas se beneficien (o padezcan) en forma injusta los atributos naturales y sociales que no les *pertenecen* en sentido estricto. Evidentemente, puede decirse que los diversos atributos naturales con los cuales he nacido me «pertenecen» en un sentido débil y contingente en cuanto a que residen accidentalmente dentro de mí, pero este sentido de propiedad o de posesión no puede determinar que yo tenga derechos especiales con respecto a estos activos o un derecho

privilegiado sobre los frutos de su ejercicio. En este sentido atenuado de la posesión, yo no soy realmente el propietario sino simplemente el guarda o depositario de los diversos atributos y activos que se ubican «aquí». Al no reconocer la arbitrariedad de la suerte, los principios de la libertad natural y de la igualdad liberal se equivocan al asumir que «mis» atributos me pertenecen en un sentido fuerte y constitutivo, con lo que permitirían que las porciones distributivas dependieran de ellos.

Expresada en términos de mérito, la objeción de Rawls a los principios de la libertad natural y de la igualdad liberal se basa en que recompensan atributos que las personas no se puede decir que merezcan. Aunque algunos pueden pensar que los afortunados merecen las cosas que los conducen a una posición de mayor beneficio,

> [...] no merecemos el lugar que tenemos en la distribución de dones naturales, como tampoco nuestra posición inicial en la sociedad. Igualmente problemático es el que merezcamos el carácter superior que nos permite hacer el esfuerzo por cultivar nuestras capacidades, ya que tal carácter depende, en buena medida, de condiciones conocidas y sociales afortunadas en la niñez, por las cuales no puede pretenderse crédito alguno. La noción de mérito no puede aplicarse aquí. (126)

Dado que nadie merece su buena suerte en la lotería genética, o un lugar inicial de privilegio en la sociedad, o incluso el carácter superior que lo motiva a cultivar sus habilidades concienzudamente, no podemos decir que nadie merezca los beneficios que producen estos atributos. Es ésta la conclusión que Nozick contradice. «No es verdad, por ejemplo, que una persona *gane Y* (el derecho de conservar una pintura que ha pintado, elogios por escribir *Teoría de la Justicia,* etc.) sólo si ha *ganado* (o de otra manera *merece)* todo lo que él usó (incluyendo dotes naturales) en el proceso de crear *Y*. Algunas de las cosas que usa, simplemente las puede tener, no ilegítimamente. No es necesario que los fundamentos que subyacen el merecimiento, sean también merecidos, *de pe a pa*» [Los subrayados son míos] (Nozick 1974: 221).

Entonces, ¿qué debemos entender de esta exposición? Si yo no tengo necesariamente que *merecer* todo lo que uso para producir una cosa a fin de merecer esa cosa, ¿de qué depende

el mérito? Nozick alega que algunas de las cosas que uso, «solamente las *tengo*, no ilegítimamente» (y, presumiblemente, es posible que esto sea arbitrario). Una vez más, entra en escena la noción de posesión. Para determinar si, el que yo tenga una cosa, no ilegítimamente, me puede capacitar para merecer lo que me ayuda a producir, debemos explorar en mayor detalle la relación entre la posesión y el mérito, y establecer una vez más a cuál sentido de la posesión se refiere.

Para este propósito, puede resultar útil considerar un reciente debate sobre la justicia y el mérito personal desarrollado por Joel Feinberg, quien analiza las bases del mérito con una admirable claridad en términos sugestivos para los argumentos a los que nos estamos dedicando (1970). Feinberg comienza con la observación de que nadie puede merecer nada a menos que exista una base para el mérito. «Un mérito sin fundamento simplemente no es mérito.» Pero surge inmediatamente la cuestión de qué *tipo* de fundamento es necesario. En las palabras de Feinberg: «No cualquier fundamento serviría». Una vez más, la noción de posesión es la que da la clave. «Si una persona merece algún tipo de tratamiento, debe, necesariamente, hacerlo en virtud de alguna actividad anterior o *característica poseída*» [Los subrayados son míos] (1970: 48).

> Una característica que me es propia no puede ser la base de un mérito tuyo a menos que de alguna manera revele alguna característica que te sea propia o se refleje en ella. En general, los hechos que constituyen la base del mérito de un sujeto deben ser hechos con respecto a ese sujeto. Si un estudiante merece una buena nota en un curso, por ejemplo, su mérito debe ser en virtud de algún hecho relacionado con él (su desempeño anterior, o sus actuales habilidades). [...] Es necesario que el mérito de una persona tenga una base, y que tal base esté integrada por algún hecho con respecto a ella. (Feinberg 1970: 58-9, 61)

El análisis que hace Feinberg, vinculando el mérito de una persona con algún hecho relacionado con ella, parece sustentar el argumento de Nozick de que «no *absolutamente todos* los fundamentos que subyacen al mérito necesitan ser merecidos en sí mismos». De hecho, la existencia de mérito en alguna característica poseída por la persona sugiere una tesis aún más fuerte que la de Nozick: que los fundamentos que subyacen al mérito *no pueden* ser merecidos en sí mismos, al igual que los fundamentos que subyacen a la posesión no pueden ser poseí-

dos. Ya hemos visto cómo la noción de la posesión requiere que en algún lugar subyacente se halle un sujeto de la posesión que no esté *en sí mismo* poseído (porque esto negaría su facultad de agencia), como sujeto «dueño de la posesión», por decirlo de alguna manera. La analogía con el mérito debe ser una *base* para el mérito, en última instancia anterior a éste. Consideremos: si el mérito presupone cierta característica poseída, y si las características poseídas presuponen algún sujeto de la posesión que no esté poseído en sí mismo, el mérito debe presuponer algún sujeto de la posesión que no esté poseído en sí mismo, y por lo tanto alguna base para el mérito que no sea merecida en sí misma. De la misma manera en que debe existir un sujeto de la posesión anterior a la posesión, también debe existir una base para el mérito anterior a éste. Es por esto por lo que la cuestión de si alguien merece (tener) este carácter férreo, por ejemplo, es notablemente difícil (ya que no resulta claro quién o qué queda para juzgar una vez que se ha eliminado su carácter), y por lo que más allá de un cierto punto, la simple pregunta general de si alguien merece ser (el tipo de) persona que es, se vuelve completamente incoherente. De alguna manera, debe haber una base para el mérito que no sea en sí merecida. Los fundamentos que subyacen al mérito no pueden ser en su totalidad merecidos en sí mismos.

Este resultado parecería confirmar ampliamente el argumento de Nozick contra Rawls de que no necesariamente tengo que *merecer* todo lo que utilizo en la producción de una cosa para merecerla, que algo de lo que utilizo puedo «sólo *tenerlo*, no ilegítimamente». Y si se puede establecer este argumento, parecería que el argumento de Rawls sobre la arbitrariedad no llega a socavar la noción de mérito después de todo. Decir, como lo hace Rawls, que yo no merezco el carácter superior que me lleva a realizar mis habilidades ya no resulta suficiente. Para negar mi mérito, es necesario demostrar que yo no *tengo* el carácter requerido, o alternativamente, que lo *tengo* pero no en el sentido requerido.

Sin embargo, éste es precisamente el argumento que su teoría de la persona le permite formular a Rawls, ya que dada su clara distinción entre el «yo» —tomado como el sujeto de la posesión puro— y los fines y atributos que posee, el «yo» queda desnudo de todo rasgo o característica substantivos que puedan considerarse como bases para el mérito. Dado el aspecto de distanciamiento de la posesión, el «yo» *en sí mismo* sufre la

pérdida de posesión. En la teoría de la persona de Rawls, el «yo», estrictamente hablando, *no tiene nada*, nada al menos en el sentido fuerte y constitutivo necesario para el mérito. En una maniobra similar a la que invoca para demostrar que el principio de diferencia no utiliza a la *persona* como un medio, sino sólo a los *atributos* de la persona, Rawls puede aceptar que cierta base no merecida es necesaria para la noción de mérito, sólo para alegar que, en una comprensión adecuada de la persona, ¡en principio esta condición nunca podría cumplirse! En la concepción de Rawls, las características que poseo no están *integradas* al «yo» sino sólo *relacionadas* con el «yo», siempre separadas por una cierta distancia. Esto es lo que las hace atributos y no componentes de la persona; son *mías* en lugar de ser *yo*, son cosas que *poseo* y no que *soy*.

Podemos ver en esta luz cómo el argumento de Rawls contra la arbitrariedad no ataca directamente al mérito, alegando que no puedo *merecer* lo que está dado arbitrariamente, sino indirectamente, alegando que no puedo *poseer* lo que está arbitrariamente dado: es decir que *yo*, en tanto sujeto de la posesión, no puedo poseerlo en el sentido constitutivo y no distanciado que es necesario para brindar un fundamento al mérito. Un bien dado arbitrariamente no puede ser un componente esencial sino sólo un atributo accidental de mi persona, ya que de otra manera mi identidad dependería de una mera contingencia, y su continuidad sería constantemente vulnerable a la transformación por la experiencia; mi categoría de agente soberano dependería de las condiciones de mi existencia en lugar de estar garantizada epistemológicamente. En la perspectiva de Rawls, no puede decirse que nadie merezca nada porque no puede decirse que nadie posea nada, al menos no en el sentido fuerte, constitutivo de la posesión, que es necesario para la noción de mérito.

Una teoría de la justicia sin mérito parecería separarse dramáticamente de las concepciones tradicionales, pero Rawls se esfuerza en demostrar que no es así. En sus primeras páginas, Rawls reconoce que «puede parecer que este enfoque no esté de acuerdo con la tradición», pero intenta asegurarnos que de hecho lo está:

> El sentido más específico que Aristóteles da a la justicia y del cual se derivan las formulaciones más conocidas, es el de abstenerse de la *pleonexia*, esto es, de obtener para uno mismo cierta

ventaja apoderándose de lo que pertenece a otro, sus propiedades, sus remuneraciones, su empleo o cosas semejantes; o negándole a una persona lo que le es debido [...] *la definición de Aristóteles presupone claramente una explicación de lo que propiamente le pertenece a una persona y de lo que le es debido. Ahora bien, creo que tales derechos se derivan muy a menudo de instituciones sociales y de las expectativas legítimas que ellas originan.* No hay razón para creer que Aristóteles hubiese estado en desacuerdo con esto y, ciertamente, tiene una concepción de la justicia social para dar cuenta de estas pretensiones. [...] No hay conflicto con la definición tradicional. [Los subrayados son míos.] (27-28)

Al comparar la justicia como imparcialidad con las concepciones tradicionales, el autor confirma su novedad en lugar de negarla. Si analizamos lo que presenta como una calificación incidental de la justicia en su concepción clásica, ésta resulta en realidad una divergencia concluyente. Tal como sugiere Rawls, las nociones tradicionales se refieren libremente a «lo que propiamente le pertenece a una persona», presumiblemente aparte de las instituciones; presuponen personas de una constitución sólida con una fijeza de carácter, algunos de cuyos rasgos se toman como totalmente esenciales. En la concepción de Rawls, sin embargo, no aparece ninguno de estos conceptos. En la medida en que una teoría de la justicia «presupone una explicación de lo que propiamente le pertenece a una persona» (en el sentido fuerte de «pertenecer»), Rawls admite efectivamente que no tiene ninguna. Y parece implicar que tampoco es razonable pensar que tal teoría de la justicia pudiera ser verdadera, dada la precedencia de la pluralidad, la prioridad de lo justo, y la teoría de la persona que requiere. Esencialmente no somos seres lo suficientemente densos como para comprender derechos y méritos antecedentes a las instituciones que los definen. Dadas estas restricciones, la única alternativa es optar por una teoría de la justicia basada en derechos que legitiman las expectativas, eliminando completamente el mérito. Rawls elude al principio este postulado, afirmando sólo que «tales derechos se derivan *muy a menudo* de instituciones sociales y de las expectativas legítimas que ellas originan» [Los subrayados son míos] (27). Pero a medida que emergen las plenas consecuencias del punto de vista de Rawls, «muy a menudo» se transforma en «siempre», ya que se vuelve claro que «tales derechos» no pueden surgir de ninguna otra manera. Aunque Aristóteles pueda no estar en desacuerdo con que los derechos

surjan de esta manera, parece bastante alejado de su perspectiva que no puedan surgir de ninguna otra manera. Al negar que la justicia tiene que ver con darle a las personas lo que merecen, la justicia como imparcialidad, después de todo, se aparta en forma decisiva de la noción tradiciona.

La aparente perspectiva de Rawls de que no se puede decir que nadie merezca nada, y su conexión con la noción del «yo» como «esencialmente no restringido», emerge en forma más clara en su argumentación sobre las expectativas legítimas y el mérito moral. Comienza reconociendo que la justicia como imparcialidad, al rechazar el mérito, se contradice con el sentido común.

> Hay una tendencia, por parte del sentido común, a suponer que la renta y la riqueza, y los bienes en general, han de ser distribuidos de acuerdo con un criterio moral. La justicia es la felicidad acorde con la virtud. A pesar de que se reconoce que este ideal nunca puede ser plenamente conseguido en su totalidad, es la concepción apropiada de la justicia distributiva, al menos en un principio *prima facie*, y la sociedad debería llevarlo a cabo en tanto las circunstancias lo permitan. Ahora bien, la justicia como imparcialidad rechaza esta concepción. Tal principio no sería elegido en la posición original. Parece no haber medio de definir el criterio requerido para esta situación. (349)

Parecería que no hay manera de definir los criterios requeridos para la virtud o el valor moral de una persona en la posición original debido a que no contamos con una teoría de la persona sustantiva antecedente a las instituciones sociales. Para que el mérito moral brindara un criterio de justicia independiente, debería sostenerse en alguna teoría sustantiva de la persona, o del valor de las personas. Pero para Rawls el valor de las personas es posterior a las instituciones, no independiente de ellas. Y por lo tanto los derechos morales de una persona deben aguardar su advenimiento.

Esto nos lleva a la distinción entre el mérito moral y las expectativas legítimas. Una vez que la persona realiza las diversas cosas establecidas que las instituciones la alientan a emprender, adquiere ciertos derechos, pero no antes. Tiene derecho a que las instituciones cumplan las demandas que anunciaron que recompensarían, pero en primer lugar no tiene derecho a que éstas se comprometan a recompensar ningún tipo de demanda en particular.

> Por lo tanto un esquema justo responde a lo que las personas están autorizadas a exigir; este esquema satisface sus legítimas expectativas basadas en las instituciones sociales. Pero lo que están autorizados a exigir no es proporcional, ni depende de su valor intrínseco. Los principios de la justicia que regulan la estructura básica, y especifican los deberes y obligaciones de los individuos, no mencionan el criterio moral, y no hay una tendencia por parte de las porciones distributivas a adaptarse a ello. (349-350)

Los principios de la justicia no mencionan el mérito moral debido a que, estrictamente hablando, no puede decirse que nadie merezca nada. En forma similar, la razón por la que los derechos de las personas no son proporcionales ni dependen de su valor intrínseco es que, en la perspectiva de Rawls, *las personas no tienen un valor intrínseco*, no en el sentido de que sea suyo antes o independientemente o aparte de lo que las instituciones justas les atribuyan.

> Lo esencial es que el concepto de valor moral no proporciona un primer principio de justicia distributiva. Esto se debe a que no puede ser introducido hasta que los principios de la justicia, y del deber y la obligación natural, han sido reconocidos. [...] [E]l concepto de valor moral es inferior a los conceptos del derecho y la justicia, y no desempeña ningún papel en la definición substantiva de las porciones distributivas. (351-352)

Rawls podría estar de acuerdo con Feinberg en que «el mérito es un concepto *moral* en el sentido de que es lógicamente anterior a las instituciones públicas y sus reglas, e independiente de ellas» (313), y por lo tanto estaría en desacuerdo con que «uno de los objetivos [de un sistema de bienes públicos] es dar a las personas lo que ellas merecen» (Feinberg 1970:86). Para Rawls, los principios de la justicia no apuntan ni a recompensar la virtud ni a dar a las personas lo que ellas merecen, sino a obtener los recursos y talentos necesarios para el servicio del interés común.

> Además, ninguno de los preceptos de la justicia aspira a una virtud compensadora. Los premios obtenidos por los talentos naturales escasos, por ejemplo, han de cubrir los costes de las enseñanzas y alentar los esfuerzos en el aprendizaje, además de dirigir las distintas capacidades hacia donde mejor se favorezca el interés común. Las porciones distributivas resultantes no se relacionan con el valor moral [...]. (349)

Para ilustrar la prioridad de las instituciones justas con respecto a la virtud y el valor moral, Rawls sugiere una analogía con la relación entre las reglas de propiedad y la ley sobre robo y hurto.

> Estos delitos, y el demérito que ocasionan, presuponen la institución de la propiedad, establecida con fines sociales prioritarios e independientes. Para una sociedad, el organizarse a sí misma con la intención de considerar el criterio moral como un primer principio, sería lo mismo que mantener la institución de la propiedad para castigar a los ladrones. El criterio: a cada uno de acuerdo con su virtud, no sería, por tanto, elegido en la posición original. (352)

La analogía es interesante, pero uno se pregunta si funciona completamente a favor de Rawls. Aunque es evidente que la institución de la propiedad tiene una *cierta* prioridad con respecto a los delitos correlativos a ella, resulta menos claro por qué la dependencia debe funcionar en una sola dirección, especialmente dado el compromiso que asume el propio Rawls con el método del equilibrio reflexivo. Por ejemplo, ¿nuestra creencia en la validez de la institución de la propiedad no es aumentada de alguna manera por una convicción de que el robo y el hurto son incorrectos? ¿Nuestra confianza en la institución de la propiedad no se vería afectada si resultara que aquellos definidos como ladrones fueran invariablemente hombres buenos y virtuosos? ¿Y qué hay de los casos más extremos? Aunque las normas y reglas que protegen la vida humana pueden indudablemente defenderse sobre una diversidad de fundamentos, tales como mantener a las personas con vida, evitar el sufrimiento, etc., ¿resulta lógicamente erróneo pensar que una de las justificaciones de las prohibiciones sobre el asesinato sería la de castigar a los asesinos?

La posición de Rawls aparece aquí como particularmente sorprendente a la luz del contraste que establece entre la justicia distributiva y la justicia retributiva, sugiriendo que en el segundo caso, alguna noción de mérito moral, después de todo, puede resultar adecuada. Es el punto de vista de que las porciones distributivas deben corresponderse con el valor moral en tanto sea posible, en palabras de Rawls, «aquellos que son castigados por violar leyes justas normalmente han cometido algún delito. Esto se debe a que el propósito de la ley penal es

defender los deberes naturales básicos [...] y los castigos tratan de conseguir este fin».

> No son simplemente un esquema de impuestos y cargas designados para poner precio a ciertas formas de conducta y, de este modo, guiar la conducta de las personas para obtener un beneficio recíproco. Sería mucho mejor si los actos prohibidos por la ley penal nunca se cometiesen. Así, la propensión a cometer tales actos es característica de un mal carácter, y en una sociedad justa los castigos legales caerán únicamente sobre aquellos que demuestren estos defectos.
>
> Es claro que la distribución de ventajas económicas y sociales es completamente diferente. Estas soluciones no se oponen, por así decirlo, a la ley criminal, así que, en tanto una castiga ciertos delitos, la otra incrementa la virtud moral. La función de las porciones distributivas desiguales es cubrir los costes de enseñanza y educación, atraer a las personas a los lugares donde más se les necesita desde un punto de vista social, etc. [...] *El considerar que la justicia distributiva y la justicia retributiva son opuestas la una a la otra es completamente erróneo, y sugiere una base moral para las porciones distributivas donde no existe.* [Los subrayados son míos.] (354)

A diferencia de los beneficios que se derivan de los acuerdos distributivos, los castigos y prohibiciones que se asocian con el derecho penal no son simplemente un sistema no moral de incentivos y desalientos diseñado para alentar algunas formas de conducta y desalentar otras. Para Rawls, las nociones morales preinstitucionales que se excluyen en la justicia distributiva de alguna manera encuentran sentido para propósitos retributivos, y existe una tendencia a que el castigo corresponda a éstos.

La duda inmediata es cómo puede corresponderse esta explicación con la analogía de la propiedad y el robo. Si la justicia retributiva se diferencia de la justicia distributiva justamente en virtud de su base moral anterior, es difícil comprender de qué manera el ejemplo de la propiedad y el robo podría demostrar la prioridad de las instituciones sociales con respecto a la virtud y al valor moral, si la prioridad sólo se aplica a la justicia distributiva. Dejando aparte esta confusión menor, la cuestión más básica es cómo puede admitir Rawls el mérito en la justicia retributiva sin contradecir la teoría del «yo» y los supuestos relacionados que la eliminaban con respecto a

los propósitos de la justicia distributiva. Si nociones tales como demandas morales preinstitucionales y valor moral intrínseco se excluyen de una teoría de la justicia distributiva en virtud de un «yo» esencialmente no restringido que es demasiado incorpóreo como para sustentarlas, es difícil percibir de qué manera la justicia retributiva podría ser diferente de una manera relevante.[11]

¿Por qué los mismos argumentos de arbitrariedad que excluyen al mérito con respecto a las porciones distributivas, no lo excluyen como fundamento del castigo? ¿La propensión a cometer crímenes es menos el resultado de factores arbitrarios desde el punto de vista moral que la propensión a hacer el bien? Y de no ser así, ¿por qué las partes en la posición original no acordarían compartir el destino mutuo con respecto a la responsabilidad penal como lo hacen con respecto a los acuerdos distributivos? Dado que bajo el velo de la ignorancia, nadie puede saber si tendrá la mala fortuna de nacer en circunstancias sociales y familiares desfavorables que lo conduzcan a una vida de crímenes, ¿por qué no podrían adoptar las partes una forma de principio de diferencia para los castigos como lo hacen con respecto a las porciones distributivas, y en efecto considerar la distribución de las responsabilidades naturales y sociales como cargas comunes?

Rawls mantiene que «aquellos que son castigados por violar leyes justas normalmente han cometido algún delito» (354). Pero supongamos que, por un acto de vandalismo, yo privo a la comunidad de cierta medida de bienestar, digamos al arrojar un ladrillo para romper una ventana. ¿Hay alguna razón para que yo tenga que soportar todos los costos negativos de mi barbarie, mientras que la persona que produjo la ventana no *merece* disfrutar de todos los beneficios de su productividad? Rawls puede responder que mi «propensión a cometer tales actos es característica de un mal carácter». Pero si la diligencia del trabajador que realizó la ventana no es una caracterís-

11. En una nota al pie, Rawls (354) cita a Feinberg aparentemente para apoyar su postulado, pero Feinberg incluye un papel para el mérito tanto en la justicia distributiva como en la retributiva. El eje de Feinberg se centra en que la justicia retributiva involucra lo que él llama el mérito polar (donde una persona merece el bien o merece el mal), mientras que la justicia distributiva involucra el mérito no polar (ya que, como en el caso de un premio, algunos lo merecen y otros no). Pero ambos casos involucran al mérito en sentido moral preinstitucional (Feinberg 1970: 62).

tica de buen carácter (en el sentido moral, preinstitucional), ¿por qué mi malicia al romper la ventana constituye una característica de mal carácter (en el sentido moral, preinstitucional)? Evidentemente, (siguiendo a Rawls, p. 126), dado un sistema justo de derecho penal, aquellos que hayan realizado lo que el sistema anuncia que penará son tratados adecuadamente en forma correlativa, y en este sentido «merecen» la pena. «[Este sentido del mérito] presupone la existencia de un esquema [retributivo] que funciona y nada tiene que ver con el problema de si el esquema mismo habrá de ser diseñado de acuerdo con el principio de diferencia o con algún otro criterio» (126).

Algunos alegan que el criminal merece su castigo en el sentido moral fuerte porque se lo merece la vileza de carácter que refleja su criminalidad. Quizás esto es lo que Rawls tiene en mente cuando escribe que «la propensión a cometer tales actos es característica de un mal carácter», y que el castigo cae adecuadamente sobre aquellos que muestran estos defectos. Debido a que el transgresor es menos valioso en este sentido, merece la mala fortuna que lo aqueja. Pero sin embargo (siguiendo a Rawls, p. 127), esta perspectiva es evidentemente errónea. El que nadie merece su lugar en la distribución de dones y defectos innatos, al igual que nadie merece su lugar inicial en la sociedad, parece ser uno de los puntos clave de nuestro razonamiento. La afirmación de que un hombre merece el carácter inferior que le impide superar sus defectos resulta igualmente problemática, dado que su carácter depende en gran medida de circunstancias familiares y sociales poco afortunadas por las cuales no puede culparse a nadie. La noción de mérito parece no aplicarse a estos casos. Pero nada de esto equivale a decir que, en términos generales, una teoría de la justicia distributiva no moral no sea compatible con una teoría del castigo moral o basada en la noción del mérito. Dadas las razones que ofrece Rawls para rechazar los acuerdos distributivos basados en el mérito, parece claramente comprometido a rechazar también los acuerdos retributivos basados en el mérito.

La aparente inconsistencia entre las teorías distributiva y retributiva de Rawls no tiene por qué afectar seriamente a la teoría como un todo. Dado el método del equilibrio reflexivo, la «justificación es cuestión del mutuo apoyo de muchas consideraciones y de que todo se ajuste conjuntamente en una vi-

sión coherente» (39). Desde el punto de vista de la teoría general de Rawls, hay poco que dependa de su teoría retributiva, aparte de la medida de plausibilidad que le otorga a la justicia como imparcialidad para aquéllos comprometidos con la idea de una noción fuerte del castigo basada en el mérito. Si la distinción que traza Rawls es acertada, no es necesario escoger entre sus intuiciones retributivas y el principio de diferencia; si no lo es, una de las dos convicciones debe ceder. Si la reflexión nos lleva a considerar que una teoría no moral del castigo es inaceptable, incluso a la luz de la arbitrariedad de las características y disposiciones criminales, el principio de diferencia (que rechaza la noción de mérito) quedaría seriamente en duda. Si, por otra parte, nuestra intuición de que los criminales merecen castigo prueba no ser más indispensable que nuestra intuición de que la virtud merece recompensa (una intuición del sentido común que Rawls rechaza explícitamente), entonces podemos ajustar nuestras intuiciones en la dirección que afirma el principio de diferencia en lugar de la que se le opone. Rechazaremos la noción de mérito como fundamento tanto de los acuerdos distributivos como de los retributivos, y así resolveremos la inconsistencia.

Pero dicha solución nos hace volver a las dificultades más profundas de una teoría de la justicia sin mérito y una noción del «yo» como esencialmente desposeído o carente de rasgos constitutivos. Nozick argumenta en contra de Rawls que los fundamentos que subyacen al mérito no necesitan ser todos merecidos en sí mismos. Pero, como hemos visto, el rechazo del mérito por parte de Rawls no depende de la tesis que refuta Nozick, sino de la noción del «yo» como un sujeto de la posesión puro, no adulterado, «esencialmente descarnado». Rawls no sostiene el punto de vista de que una persona sólo puede merecer una cosa que produce si merece todo lo que utilizó en su producción, sino el de que nadie posee nada en el sentido fuerte y constitutivo necesario como fundamento del mérito. No podemos decir que nadie merezca nada (en el sentido fuerte, preinstitucional) porque no podemos decir que nadie posea nada (en el sentido fuerte, constitutivo). Ésta es la fuerza filosófica detrás del argumento basado en la arbitrariedad.

El hecho de que el argumento basado en la arbitrariedad funcione de esta manera puede verse reflejado en la transición desde la libertad natural a la igualdad de oportunidades para

la concepción democrática, tal como la traza Rawls, como etapas en la desposesión de la persona. Con cada transición, un «yo» sustantivo, «lleno [...] de rasgos particulares», se va despojando progresivamente de características que una vez se consideraron esenciales para su identidad; a medida que cada vez más rasgos se consideran arbitrariamente dados, dejan de concebirse como componentes presuntos y se ven relegados a meros atributos del «yo». Más se transforma en *mío,* y menos sigue siendo *yo*, para volver a nuestra anterior formulación, hasta que en última instancia el «yo» se ve completamente purgado de componentes empíricos y transformado en una condición de la agencia que se encuentra más allá de los objetos de su posesión. Podemos reconstruir la lógica del argumento de Rawls de la manera siguiente.

En un extremo del espectro, incluso antes de que aparezca la libertad natural, se encuentran las sociedades aristocráticas y de castas: en tales sociedades, las perspectivas de vida de una persona están vinculadas a la jerarquía dentro de la cual nace y de la cual su persona resulta inseparable. En este caso, el «yo» está muy condicionado, casi confundido con su condición, inmerso en su situación. El sistema de la libertad natural elimina la condición fija del nacimiento como un componente supuesto de la persona, y considera a cada una como libre, dadas sus capacidades y recursos, para competir en un mercado lo mejor que pueda, y recoger su recompensa. Cuando el contrato reemplaza al rango como la base de las expectativas, el sistema de libertad natural repara la arbitrariedad de las sociedades jerárquicas tomando una concepción más estrecha de la persona, por decirlo así, al considerarla distinta y separable de su entorno. De todas maneras, continúa dándose cierta arbitrariedad, especialmente en la forma de las contingencias sociales y culturales. En el régimen de libertad natural, las perspectivas de vida de una persona están regidas por factores que ya no pueden adscribirse a la persona (en el sentido fuerte, constitutivo), al igual que su rango hereditario. Habiendo liberado a la persona de su bagaje jerárquico, el principio de libertad natural aún tiene una concepción del «yo» sólidamente constituido, cargado por los accidentes de las contingencias sociales y culturales. Y es por esto que se avanza hacia el concepto de libertad de oportunidades, que libera al «yo» de sus accidentes sociales y culturales al igual que de su rango hereditario. En una «meritocracia justa», los efectos de

la clase y las desventajas culturales se consideran como un reflejo de la sociedad más que de la persona. Aquellos con talentos comparables y «[...] la misma voluntad de utilizarlos, deben tener las mismas perspectivas de éxito, independientemente de su posición inicial en el sistema social, es decir, de la clase dentro de la cual han nacido» (94). De esta manera la concepción meritocrática extiende la lógica de la libertad natural al adscribir menos al «yo» y más a su situación.

Pero incluso el principio de igualdad de oportunidades, al recompensar el esfuerzo individual, tiene una concepción demasiado expansiva del «yo». Ya que «el esfuerzo que una persona desea hacer está influido por sus capacidades naturales, sus conocimientos y las alternativas que se le ofrecen. Los mejor dotados tratan de hacer un esfuerzo consciente, y no parece haber medio de desestimar su mejor fortuna» (351). El «yo» se encuentra todavía sobrecargado. Dada su arbitrariedad, incluso el carácter que determina la motivación de una persona no puede considerarse adecuadamente como un componente esencial de su identidad. Y así finalmente avanzamos hacia la concepción democrática, en la que el «yo», que ha perdido todos sus atributos aleatoriamente dados, asume una especie de jerarquía supraempírica: esencialmente descarnado, anticipadamente limitado y dado con prioridad a sus fines, un sujeto puro de agencia y de posesión, finalmente débil. No sólo mi carácter sino incluso mis valores y convicciones más profundas están relegadas a lo contingente, como rasgos de mi condición y no como componentes de mi persona. «El que tengamos una concepción del bien en lugar de otra no es relevante desde un punto de vista moral. Al adquirirla estamos influidos por el mismo tipo de contingencias que nos lleva a eliminar nuestro conocimiento de nuestro sexo y nuestra clase» (Rawls, 1975: 537). Sólo de esta manera es posible instalar al «yo» como invulnerable, asegurar su soberanía de una vez y para siempre en un mundo que amenaza permanentemente con engullirlo. Sólo si la suerte del «yo» se separa de esta manera de la suerte de sus atributos y objetivos, sujetos como están a los vaivenes de las circunstancias, podemos preservar su prioridad y garantizar su facultad de agencia.

Ésta es la visión de la persona que Nozick y Bell, como defensores de la libertad natural y de la meritocracia respecti vamente, rechazan enfáticamente, incluso aunque no son realmente específicos acerca de la concepción del «yo» en la que se

basan. Ambos objetan que el argumento basado en la arbitrariedad, aplicado en forma consistente, nos lleva inevitablemente a la disolución de la persona, y a la abdicación de la responsabilidad individual y de la elección moral. «Esta línea argumentativa sólo puede lograr bloquear la introducción de las elecciones y actividades autónomas de la persona (y sus resultados) recurriendo a atribuir *todo* lo que es de valor en una persona enteramente a cierto lapso de factores "externos"», afirma Nozick. Remitiéndonos a su argumento contra la noción del acervo común, Nozick cuestiona si en la explicación de Rawls podemos encontrar alguna concepción coherente de la persona, y de ser así, si continúa siendo una clase de persona que valga toda la tinta vertida por el liberalismo deontológico en ocuparse de ella.

> De esta manera, denigrar la autonomía de una persona y la genuina responsabilidad de sus acciones es una línea arriesgada de tomar para una teoría que, de otra manera, desea reforzar la dignidad y el respeto propio de seres autónomos; especialmente para una teoría que se funda tanto (incluyendo una teoría del bien) sobre las opciones de las personas. Podemos dudar que el cuadro no exaltado de los seres humanos que la teoría de Rawls presupone, y sobre la cual descansa, pueda encajar conjuntamente con la idea de dignidad humana que la teoría, para eso formulada, quiere alcanzar y abrazar. (Nozick 1974: 211)

Bell resume su objeción en un epigrama: «La persona ha desaparecido. Sólo quedan los atributos» (Bell 1973: 419). Mientras que Rawls intenta asegurar la autonomía del «yo» desprendiéndolo del mundo, sus críticos alegan que termina disolviendo al «yo» al intentar preservarlo.

Para recapitular nuestra versión reconstruida del debate entre Rawls y Nozick sobre el tema del mérito: Nozick alega en primer lugar que la arbitrariedad de los atributos no afecta negativamente al mérito, ya que el mérito puede depender no sólo de las cosas que merezco, sino también de las cosas que *tengo*, no ilegítimamente. La respuesta de Rawls es invocar la distinción entre el «yo» y sus posesiones en la versión más fuerte de esa distinción, y afirmar que, estrictamente hablando, *no hay nada* que yo, en tanto sujeto de la posesión, *tenga* —nada que esté anexado o vinculado con *mi persona*— al menos nada en el sentido fuerte y constitutivo de la posesión que es necesario como fundamento del mérito. La réplica de No-

zick se basa en que su defensa es limitada, porque tiene la consecuencia de dejarnos con un sujeto *tan* desposeído de características empíricamente identificables que se parece nuevamente al sujeto trascendente o incorpóreo kantiano que Rawls se proponía evitar. El individuo se vuelve inviolable sólo al volverse invisible, y esto pone en duda la dignidad y autonomía que este liberalismo intenta asegurar por encima de todo.

Demandas individuales y demandas sociales: ¿qué pertenece a quién?

Pero Nozick tiene otra objeción, independiente de la primera, que se formula aproximadamente así: aunque Rawls esté en lo correcto en cuanto a que la arbitrariedad socava la posesión individual y por lo tanto el mérito individual, el principio de diferencia no es su resultado inevitable. Aún podría ser verdadera una versión del principio de la libertad natural (aquí Nozick prefiere hablar en términos de su propia versión, a la que denomina «teoría de los derechos»). Porque, aunque nadie mereciera ninguno de sus atributos naturales, podría sin embargo darse que las personas tuvieran derechos aún sobre ellos, y sobre lo que se pueda derivar de ellos.

En el mejor de los casos, según este punto de vista, la defensa de Rawls del principio de diferencia se ve socavada por el argumento basado en la arbitrariedad. Demostrar que los individuos, como tales, no merecen o no poseen «sus» atributos no equivale necesariamente a demostrar que la sociedad como totalidad *sí* los merece o los posee. Simplemente porque los atributos accidentalmente ubicados en mí no sean *mis* atributos no implica necesariamente, como Rawls parece pensar, que sean bienes comunes, en lugar de ser los bienes de ninguno. Si no puede decirse que me pertenecen propiamente a mí, ¿por qué asumir automáticamente que pertenecen a la comunidad? ¿El hecho de que se ubiquen dentro de la *comunidad* es menos accidental, o menos arbitrario desde un punto de vista moral? Y de no ser así, ¿por qué no considerarlos como bienes de libre circulación, no vinculados por anticipado con *ningún* sujeto de la posesión, ya sea individual o social?

Es necesario ser más preciso sobre los términos de la relación entre la persona y los bienes que presenta. Tres descripciones parecen cubrir los casos posibles: dependiendo del sentido de la posesión al que nos refiramos, se me puede describir como el propietario, como el guardián o como el depositario de los bienes que poseo. En la versión más fuerte, la noción de propiedad implica que tengo derechos absolutos, exclusivos y no calificados con respecto a mis recursos; en su versión más moderada, que tengo ciertos derechos privilegiados con respecto a ellos, una serie de derechos que, aunque no son ilimitados, son al menos más amplios con respecto a mis atributos que cualquier serie de derechos que pueda alegar cualquier otra persona sobre ellos. Éste es el sentido de la posesión individual contra la cual se dirige el argumento basado en la arbitrariedad de Rawls y al que desvirtuaría de tener éxito.

Si Rawls tiene razón, y no se me puede describir apropiadamente como el propietario de mis activos y atributos, subsisten dos explicaciones alternativas. Una es la noción de tenencia, que niega la propiedad individual a favor de otro dueño o sujeto de la posesión final del cual la persona individual es un agente. Decir que soy el custodio de los bienes que presento significa implicar que los posee otro sujeto, en cuya representación o en cuyo nombre, o por cuya gracia los cultivo y ejercito. Ésta es una noción de posesión que recuerda la noción de propiedad del primer cristianismo, en el cual el hombre tenía lo que tenía como guardián de los bienes que en realidad pertenecían a Dios, y es esta noción la que se adecua a las distintas nociones comunitarias de la propiedad.[12]

12. El reverendo Vernon Bartlet (1915: 87-98) afirma que según «la actitud cristiana esencial» frente a los derechos de propiedad de cualquier individuo eran «puramente relativos, no sólo frente a los derechos absolutos de Dios como *productor* y *poseedor* de todas las cosas y de todas las personas, sino también en comparación con los derechos derivativos o humanos supremos de la sociedad como representante de un bienestar común, del cual el del individuo es sólo una parte dependiente, que debería estar limitada por los derechos de todos a la condición de bienestar personal. [...] Los principios prácticos resultantes, frente a la subordinación de la propiedad por parte de Dios y de la sociedad, y la obligación moral de fidelidad en esta relación como condición de cualquier derecho correlativo de disfrute privado personal, están demasiado arraigados en las enseñanzas de Cristo y sus parábolas, como para necesitar ser demostradas». Bartlet cita a san Pablo, I Corintios IV, 7: «Pues, quién te distingue a ti? ¿Qué tienes que no hayas recibido? Y, si lo has recibido, ¿por qué te glorías como si no lo hubieses recibido?

En el tercer caso, yo no soy ni el propietario ni el guardián, sino el depositario de los activos y atributos que se ubican accidentalmente en mi persona. La noción del individuo como el depositario de recursos no presupone otros sujetos de posesión a quien éstos pertenecen en última instancia, pero descarta completamente la noción de posesión. En tanto soy el depositario de atributos naturales, no es necesario plantearse a quién *pertenecen* en última instancia esos bienes: el que residan en mí no tiene consecuencias con respecto a las demandas que yo u otras personas puedan hacer sobre ellos.

Ahora bien, en términos de estas distinciones, Nozick alega en efecto que el argumento de Rawls basado en la arbitrariedad, incluso aunque tuviera éxito en desvirtuar a la posesión individual y por lo tanto al mérito, sólo implica la tercera descripción —que los individuos son depositarios de atributos— y no la segunda. Pero si esto es todo lo que el argumento basado en la arbitrariedad establece, no conduce al principio de diferencia, ya que este principio debe presuponer la segunda descripción, que afirma que yo soy el guardián de los bienes con respecto a los cuales la comunidad en su totalidad tiene un derecho anterior. Si todo lo que la arbitrariedad significa es que yo soy el depositario de bienes que no pertenecen a nadie en particular, no podemos asumir que la comunidad los posea más que yo mismo. Sería como si los atributos, y los beneficios que se derivan de ellos, cayeran como el maná del cielo (la imagen fue acuñada por Nozick), sin marcas de derechos anteriores y sin estar vinculados a ningún sujeto de la posesión ya sea individual o social.

¿Cómo deberían ser distribuidos entonces? ¿Sobre qué base se evaluarían las demandas conflictivas sobre un botín de generación espontánea como éste? Desde el punto de vista del mérito, parecería que no hay fundamentos sobre los cuales decidir entre dejar los atributos donde están o tratar de distribuirlos de alguna otra manera. A menos que la distribución acorde con tales supuestos se considere como una cuestión de indiferencia moral, los derechos de las personas deben depender de consideraciones distintas a las nociones de posesión o de mérito. Rawls parece dispuesto a coincidir en este punto. ¿Pero cuáles podrían ser estas consideraciones? Es en este punto donde se separan: cada uno piensa que puede aducir consideraciones no relacionadas con el mérito para superar la indiferencia a favor de su propia concepción.

Para Nozick, la ausencia de mérito crea una presunción a favor de permitir que los atributos queden donde están, al menos una vez que se acepta que las cosas no llegan a este mundo como el maná del cielo, sino que ingresan en éste como pertenencias de alguien, vinculadas a personas particulares.

> Pero como las cosas surgen ya poseídas (o con acuerdos ya celebrados sobre cómo deben ser poseídas), no hay ninguna necesidad de buscar alguna pauta que satisfacer para pertenencias no poseídas [...] La situación no es la apropiada para preguntarse: «¿Después de todo, qué va a ser de estas cosas?, ¿qué vamos a hacer con ellas?». En el mundo donde no hay maná del cielo, en el que las cosas se tienen que hacer o producir o transformar por las personas, no hay ningún proceso separado de distribución para que una teoría de la distribución sea una teoría de este proceso separado. (Nozick 1974: 216)

Nozick argumenta que si una persona tiene un atributo al cual ninguna otra persona tiene derecho, entonces, aunque es posible que no *merezca* tal atributo, tiene sin embargo *derecho* sobre él, y a lo que se derive de él a través de procesos que no violen los derechos de ninguna otra persona. «Ya sea que los recursos naturales de las personas sean arbitrarios o no, desde un punto de vista moral, las personas tienen derecho a ellos, y a lo que resulte de ellos» (Nozick 1974: 222).

Para Rawls, por otra parte, la ausencia de mérito individual crea una presunción a favor de considerar la distribución de talentos como un acervo común. La falta de mérito o de un concepto preinstitucional de la virtud significa que las instituciones no están limitadas por demandas morales antecedentes en su búsqueda de la virtud primaria de la justicia social. En este sentido, la analogía del maná del cielo es adecuada. La cantidad de atributos que otorga la suerte no es justa ni injusta. «Éstos son hechos meramente naturales. Lo que puede ser justo o injusto es el modo en que las instituciones actúan respecto a estos hechos» (125). No existe razón para permitir que los atributos y los beneficios que se derivan de ellos permanezcan donde están. Esto sería simplemente incorporar y afirmar la arbitrariedad de la naturaleza. El descubrimiento de que la virtud y los derechos son posteriores a las instituciones sociales, en lugar de restricciones, constituye un motivo para continuar la búsqueda de la justicia aún con mayor denuedo, y no para congelar la arbitrariedad tal como se verifica.

¿Qué debemos entender, entonces, de los intentos de superar la aparente indiferencia moral creada por la presunción de la ausencia de mérito? Nozick está dispuesto a aceptar que las personas pueden no merecer sus atributos naturales, pero alega que tienen derecho a ellos de todas maneras. «Si las personas tienen *X*, y el que tengan *X* (independientemente de si merecen tenerlo o no) *no* viola el derecho (en el sentido de Locke) o retribución sobre *X* de nadie más, e *Y* resulta (surge, etc.) por un proceso que no viola derechos (en el sentido de Locke) o retribuciones de ningún otro, entonces la persona tiene derecho a *Y*» (Nozick 1974: 221). Pero no demuestra cuál es el fundamento de esto, ni resulta claro cuál es precisamente la diferencia entre mérito y derecho. Rawls y Feinberg concuerdan en que «el mérito es un concepto *moral* en el sentido de que es lógicamente anterior a las instituciones públicas y sus reglas, e independiente de ellas» (Feinberg 1970: 87). Los derechos, por el contrario, son demandas que pueden surgir únicamente bajo las reglas o condiciones de calificación de instituciones ya establecidas, lo que Rawls describe como expectativas legítimas fundadas en instituciones sociales. Se desprende de este punto de vista que, a fin de diseñar o evaluar instituciones sociales, los derechos de las personas, en tanto *derivados* de las instituciones, carecen de toda fuerza moral o crítica. Evaluar la justicia de una institución a la luz de los derechos que atribuye a las personas sería como juzgar la validez de una regla a la luz de las demandas a las que da lugar; o volviendo a nuestro argumento anterior, sería apelar a una medida de evaluación completamente implicada en el objeto de la evaluación. Es por esta razón que el concepto de derecho no puede brindar un primer principio de la justicia. Tal como explica Rawls, «presupone la existencia de un esquema cooperativo que funciona y nada tiene que ver con el problema de si el esquema mismo habrá de ser diseñado de acuerdo con el principio de diferencia o con algún otro criterio» (126).

Nozick nunca llega a reconciliarse con esta dificultad. Al formular su argumento, adopta explícitamente el lenguaje de los derechos y no el del mérito, pero no reconoce su menor fuerza moral. En el uso que le da Nozick, el concepto de derecho funciona como el mérito, pero sin que se establezcan en ningún momento sus credenciales preinstitucionales. Comienza con la premisa de que: «Las personas tienen derecho a sus dotes naturales» (Nozick 1974: 222). Pero nunca explica por

qué las personas tienen derecho a sus atributos en algún sentido de derecho lo suficientemente fuerte como para respaldar el argumento.

En cierto momento, Nozick parece alegar que las personas tienen derecho a las cosas que «sólo *tienen,* no ilegítimamente», presumiendo que sus atributos naturales se encuentran entre ellas. Pero el tener algo, en forma que no sea ilegítima, no es lo mismo que tener derecho a ello; es simplemente tenerlo en un sentido no especificado de la posesión. Aunque mi inteligencia, o fuerza física, o buena salud pueden estar entre las cosas que tengo no ilegítimamente, no se deduce que tengo derecho a esas cosas, ya que los derechos dependen, como ya hemos visto, de cierto esquema de cooperación que ya funciona.

En otros momentos, Nozick parece argumentar que las personas tienen derecho a sus atributos naturales y a los beneficios que se derivan de ellos en un sentido de «derecho» anterior a las instituciones sociales, lo que equivale a decir en un sentido de derecho equiparable al mérito. («No es necesario que los fundamentos que subyacen el merecimiento, sean también merecidas, *de pe a pa*»). Esto exigiría que poseyera los atributos relevantes en el sentido fuerte de la posesión adecuado al fundamento del mérito, y a su vez exige una teoría de la persona dentro de la cual yo posea algunas cosas, por lo menos, como componentes y no meramente como atributos del «yo». Pero la teoría de la persona de Nozick no es fácil de discernir. Ataca el concepto del yo «purificado» implícito en la teoría de Rawls en tanto que parece radicalmente contrapuesto a nuestra noción más conocida de nosotros mismos como seres «llenos [...] de rasgos particulares» (224), y luego objeta que un «yo» así «purificado», aunque fuera coherente, parece poner en peligro no sólo el mérito individual, sino también nociones tan indispensables como la autonomía individual, la responsabilidad, e incluso la dignidad humana y el respeto por sí mismo, que el liberalismo deontológico se propone afirmar (Nozick 1974:211). Se trata sin lugar a dudas de objeciones poderosas. Pero poco contribuyen a articular una teoría de la persona sustantiva que logre conservar una persona «llena [...] de rasgos particulares», al mismo tiempo que evita las dificultades morales y epistemológicas. Rawls se identifica con un «yo» cargado con contingencias. Una cosa es simplemente afirmar lo que en cierto sentido es innegable, que estamos «llenos [...] de rasgos particulares», y otra demostrar que esto

es verdadero de una manera que no esté sujeta a las incoherencias contrapuestas asociadas con un «yo» radicalmente situado, indefinidamente condicionado por su entorno y constantemente sujeto a la transformación por la experiencia. Por esta razón, la solución propuesta por Nozick a la indiferencia moral resulta únicamente de interés crítico.

Con respecto a la propuesta de Rawls, si Nozick no logra demostrar que la ausencia del mérito individual conduce a una pretensión de dejar los atributos naturales donde están, ¿qué grado de éxito logra Rawls en demostrar una presunción a favor de un derecho social general a tales bienes? De acuerdo con Nozick, la perspectiva de Rawls es que «todos tienen algún derecho o reclamo sobre la totalidad de dotes naturales (reunidas), sin que nadie tenga reclamos diferentes» (Nozick 1974: 224), y de hecho la noción del acervo común parece implicar tal punto de vista. La cuestión es qué tipo de «derecho o reclamo» está en juego, y cómo puede establecerse. Surgen dos posibilidades: la demanda general social a la totalidad de los atributos naturales puede ser una demanda basada en el mérito, o puede ser una demanda de un derecho, o una expectativa legítima. Si afirmamos la demanda basada en el mérito, tal demanda es por lo menos socavada por la teoría de Rawls, ya que como hemos visto, el argumento basado en la arbitrariedad funciona solamente para desvirtuar el mérito individual, no necesariamente para instalar un mérito social. Además, para que la comunidad como un todo merezca los atributos naturales que caen dentro de su ámbito y los beneficios que se deriven de ellos, es necesario asumir que la sociedad tiene algún tipo de rango preinstitucional del cual los individuos carecen, ya que sólo de esta manera se podría decir que la comunidad posee sus atributos en el sentido fuerte, constitutivo de la posesión y necesario como fundamento del mérito. Pero este punto de vista contradiría los supuestos individualistas de Rawls, en especial su opinión de que la sociedad no constituye «un todo orgánico con una vida propia, distinta y superior a la de sus miembros en sus relaciones con los demás» (302).

La alternativa sería considerar la demanda de la sociedad sobre la distribución de los bienes como un derecho basado en una expectativa legítima fundada en instituciones sociales «establecidas con fines sociales prioritarios e independientes» (352), es decir, prioritarios e independientes del derecho en sí. Rawls puede argumentar en esta línea que el principio de di-

ferencia no supone que la sociedad tenga una categoría preinstitucional que se le niega a los individuos, sólo que las partes en la posición original acordarían considerar la distribución de talentos naturales como un acervo común y compartir los beneficios de esta distribución. La noción de sociedad como la propietaria de los atributos naturales de los cuales los seres humanos son guardianes sería considerada entonces el resultado de la posición original y no una de sus premisas.

Pero esto nos deja en la duda con respecto a «los fines sociales prioritarios e independientes», y la fuente de su prioridad e independencia. De alguna manera el derecho de la sociedad sobre la distribución de activos debe ser relevante antes de que se establezcan las instituciones. *Algunas* demandas sociales sobre la distribución de bienes parecerían un presupuesto necesario del acuerdo en la posición original, de otra forma las partes estarían deliberando sobre cómo distribuir porciones que no son (aún) legítimamente *suyas*. Dentro de esta idea general, Nozick argumenta

> ¿Se preguntarán alguna vez las personas en la posición original si *ellas* tienen el *derecho* a decidir cómo debe dividirse todo? Quizás razonan que puesto que están decidiendo esta cuestión, están obligados a suponer que tienen derecho a hacerlo. De esta forma, personas particulares no pueden poseer derechos particulares sobre pertenencias puesto que entonces no tendrían el derecho de decidir juntas cómo deben dividirse todas las pertenencias. (Nozick 1974: 197n)

Rawls podría responder a esto afirmando que no hay una demanda social antecedente implicada, ya que las partes en la posición original no se enfrentan, estrictamente hablando, con la cuestión moral de cómo deberían asignar las parcelas distributivas (que implicaría de hecho tener un «derecho» anterior a decidir), sino sólo con la cuestión de juicio de cómo, dadas las restricciones relevantes sobre la información, etc., habrían de *preferir* que se repartieran las porciones individuales desde el punto de vista del interés individual. Rawls podría recordarnos que la posición original no es un lugar real de distribución, sino una forma de pensamiento.

De todas maneras deseamos saber por qué *esta* forma de pensamiento es adecuada a la cuestión de la justicia distributiva, y si la adecuación de esta forma de pensamiento no de-

pende en sí de que las partes en la posición original tengan un derecho antecedente sobre el total de los atributos naturales, ya que incluso aunque las partes en la posición original razonaran, por una cuestión de prudencia racional, de la manera en que Rawls alega que lo harían, no resulta inmediatamente claro, en ausencia de un derecho social anterior, por qué su elección colectiva debería determinar la distribución justa de estos atributos. Esto nos lleva al tema de la justificación y a la pregunta de cómo sirve el argumento de la posición original para justificar los principios que surgen como resultado. Si el derecho de la sociedad a la distribución del acervo natural pudiera demostrarse como un producto del acuerdo original en lugar de su premisa, Rawls habría resuelto la indiferencia a favor del principio de diferencia sin tener que atribuir a la totalidad de la sociedad un derecho antecedente basado en el mérito. Por otra parte, si la noción de acervo común resultara un presupuesto del acuerdo original, Rawls podría superar la indiferencia solamente basándose en forma implícita en un postulado de mérito social, y por lo tanto basándose en que un sujeto de la posesión más amplio, presumiblemente la comunidad, poseería los propios atributos que llevamos individualmente. Una vez más una dimensión intersubjetiva irrumpiría en el proyecto individualista de Rawls. Aunque tal resultado causaría considerable daño a la ética deontológica que Rawls intenta defender, trataré de demostrar que es inevitable en su versión de la teoría del contrato.

III

Teoría contractualista y justificación

La evaluación de la validez del principio de diferencia nos conduce en última instancia a la cuestión de la justificación, y en particular a la cuestión de por qué la posición original ofrece una manera adecuada de pensamiento sobre la justicia. Aunque asumiéramos hipotéticamente que las partes en la posición original elegirían realmente los principios que Rawls afirma que elegirían, ¿por qué es *ésta* una razón para creer que esos principios son justos? Rawls postula que los principios de la justicia son aquellos que «las personas libres y racionales interesadas en promover sus propios intereses aceptarían en una posición inicial de igualdad» (28), y que la posición original es «el *status quo* inicial apropiado que asegura que los acuerdos fundamentales alcanzados en ella sean imparciales» (35). «Así entendida, la cuestión de la justificación se resuelve elaborando un problema de deliberación: tenemos que averiguar qué principios sería racional adoptar dada la situación contractual. Esto conecta la teoría de la justicia con la teoría de la elección racional» (35).

Pero no resulta claro de qué manera la posición original confiere una condición moral a los resultados de un ejercicio de elección racional, ni resulta obvio en qué consiste la base de la justificación del argumento de la posición original. La cuestión de la justificación se ve complicada por el hecho de que Rawls parece apoyarse simultáneamente en dos formas diferentes de justificación; la primera apela al método del equilibrio reflexivo, y la segunda a la tradición del contrato social.

Separar los papeles que ambas juegan plantea ciertas dificultades (Lyons, *Reading Rawls,* 141/68).

Por el momento, sin embargo, dejaremos a un lado estas dificultades a fin de concentrar nuestra atención en el aspecto contractualista de la justicia como imparcialidad. En tanto los principios de la justicia dependen para su justificación de una apelación contractualista, ¿en qué consiste la fuerza moral de este argumento? La exploración de esta cuestión debería brindarnos un nuevo control de la coherencia interna de la concepción de Rawls, y de la articulación entre la teoría de la justicia y su teoría de la persona relacionada. Trataré de demostrar que el argumento basado en la posición original puede verse como una justificación de sus resultados solamente a un coste considerable para ciertos supuestos voluntaristas e individualistas que resultan centrales para el proyecto deontológico.

La moralidad del contrato

Rawls ubica su teoría de la justicia en la tradición de la teoría del contrato social que se remonta a Locke, Rousseau y Kant. La «idea directriz» es que los principios de la justicia son el objeto de un acuerdo original. «Así pues hemos de imaginarnos que aquellos que se entregan a la cooperación social eligen, en un acto conjunto, los principios que han de asignar los derechos y deberes básicos y determinar la división de los beneficios sociales» (28). En el diseño de un contrato social «un grupo de personas tiene que decidir de una vez y para siempre lo que para ellas significará justo o injusto», y los principios que eligen «habrán de regular toda la crítica y reforma subsecuente de las instituciones» (29-30). En este sentido, el contrato original parecería una especie de contrato jurídico.

Pero para Rawls, como para algunos de sus predecesores contractualistas, el acuerdo original no se trata de un verdadero contrato histórico, sino de uno hipotético (30). Su validez no depende de términos que en realidad se hayan acordado, sino de la idea de que *hubieran* sido acordados bajo las condiciones hipotéticas requeridas. De hecho, el contrato social hi-

potético de Rawls es incluso más ficticio que la mayoría. No solamente su contrato nunca se dio en realidad; se imagina que tiene lugar entre una clase de seres que nunca existieron, seres afectados por el tipo de compleja amnesia necesaria para el velo de la ignorancia. En este sentido, la teoría de Rawls es doblemente hipotética. Imagina un hecho que nunca se realizó realmente con respecto a una clase de seres que nunca existieron en realidad.

Pero esto parecería afectar a la analogía moral que aporta mucha de su fuerza intuitiva a la teoría del contrato. Una vez que el contrato social se vuelve hipotético, el acuerdo original ya no es un contrato jurídico, sino sólo un contrato que *podría* haber sido jurídico pero que nunca lo fue. Y como afirmó Ronald Dworkin: «Un contrato hipotético no es un pálido reflejo de un contrato real; no es en absoluto un contrato» (1977ª: 17-18). ¿Cómo puede entonces servir para justificar los principios que se derivan de él, o para certificar su categoría de principios de la *justicia*?

Para responder a esta pregunta, podemos comenzar por abordar una cuestión más sencilla, e investigar la fuerza moral de los contratos y acuerdos en general. Una vez que lleguemos a una conclusión sobre cómo funciona la justificación en los contratos verdaderos, podremos ver más claramente cómo funciona en el caso de los hipotéticos.

Cuando dos personas llegan a un acuerdo podemos típicamente evaluar su justicia desde dos puntos de vista. Podemos preguntarnos sobre las condiciones en las que se realizó el acuerdo, o sea, si las partes obraron en libertad o bajo coacción, o podemos preguntarnos sobre los términos del acuerdo, es decir, si ambas partes recibieron una porción justa. Aunque estas dos consideraciones pueden bien estar relacionadas, no son en absoluto idénticas, y excluyendo algún supuesto filosófico especial que consideraremos más tarde, no pueden normalmente ser reducidas una a la otra. En términos prácticos, es más probable que un contrato acordado libremente contenga términos justos, y un intercambio (substancialmente) justo puede bien sugerir un contrato libre y no uno donde hubo coerción, pero no hay una conexión necesaria en ningún sentido.

De cualquier acuerdo contractual, no importa cuán libre sea éste, siempre es inteligible y a menudo razonable formular la siguiente pregunta: «¿Pero es *justo* lo que se acordó?», y

esta pregunta no puede traducirse meramente como la pregunta tautológica: «¿Pero lo que se acordó es lo que se acordó?». Lo que lo *hace* justo no es solamente que haya sido acordado, sino que ésta es otra cuestión.

En forma similar, una transacción o acuerdo, no importa cuán justa sea, está abierta en principio a la siguiente pregunta: «¿Pero fue acordado libremente, este acuerdo justo?», y esta pregunta no puede reducirse a la pregunta tautológica: «¿Pero este acuerdo justo es un acuerdo justo?» Lo que *hace* libre a una transacción no es que haya terminado en forma justa; ser tratados en forma justa ni nos hace libres ni implica que seamos libres. Ésta, también, es otra cuestión.

La distinción entre estos dos tipos de preguntas sugiere que podemos concebir a la moralidad de un contrato como consistente en dos ideales relacionados pero separables. Uno es el ideal de la autonomía, que percibe al contrato como un contrato entre voluntades, cuya moralidad consiste en el carácter voluntario de la transacción. El otro es el ideal de reciprocidad, que ve al contrato como un instrumento de beneficio mutuo, cuya moralidad depende de la justicia subyacente del intercambio.[13]

Cada ideal sugiere una base diferente para la obligación contractual. Desde el punto de vista de la autonomía, la fuerza moral del contrato se deriva del hecho de su acuerdo voluntario; cuando concluyo libremente un acuerdo, estoy obligado por sus términos, cualesquiera sean éstos. Ya sea que sus cláusulas sean justas o injustas, favorables o severas, yo «me las busqué», y el hecho de que sean impuestas voluntariamente brinda al menos una razón por la que estoy obligado a cumplirlas.

El ideal de reciprocidad, por otra parte, deriva la obligación contractual de los beneficios mutuos de los acuerdos cooperativos. Mientras que la autonomía apunta al contrato mismo como la fuente de las obligaciones, la reciprocidad se concentra *a través* del contrato en un requisito dado por el trasfondo moral de cumplir con los acuerdos justos, y por lo tanto implica un principio moral independiente por el cual puede evaluarse la justicia de un intercambio. En el caso de la reciprocidad, el énfasis no está tanto en el hecho del contrato como en

13. En las obras de Fried (1981) y Atiyah (1979) se encontrarán argumentos valiosos sobre los fundamentos morales del derecho contractual, que enfatizan los ideales de autonomía y reciprocidad respectivamente.

los beneficios de los que disfruto; los contratos son obligatorios no porque se concluyen libremente sino porque (o en tanto que) tienden a producir resultados que son justos.

En esta explicación de la obligación, podemos ver que cada ideal destaca lo moralmente incompleto del otro. Desde el punto de vista de la autonomía, mis obligaciones están limitadas a aquéllas que contraigo voluntariamente, pero esto puede incluir cláusulas onerosas o demasiado severas. Desde el punto de vista de la reciprocidad, los acuerdos rigurosos tienen menor fuerza obligatoria, pero por otra parte, se desdibuja la necesidad del consentimiento, y puedo estar obligado en virtud de beneficios que no deseo o dependencias más allá de mi control. En el primer caso puedo estar obligado por términos que resultan injustos; en el segundo puedo estarlo en maneras que no elegí.

Finalmente, cada explicación de las obligaciones contractuales relaciona al contrato con su justificación de maneras diferentes. Dentro del ideal de la autonomía, el contrato *imparte* la justificación; en tanto sea libre, el proceso mismo sirve para justificar el resultado, «cuaquiera que sea este». Dentro del ideal de reciprocidad, por el contrario, el contrato *se aproxima* a la justicia en lugar de *conferirla*; el proceso es instrumental para un resultado justo, y no definitorio de éste. En el primer caso, un resultado justo se define como derivado de un proceso que es libre; en el segundo, un proceso libre es simplemente un medio de llegar a un resultado que es (independientemente) justo.

A diferencia de las obligaciones contraídas voluntariamente, las obligaciones que surgen del ideal de reciprocidad deben presuponer algún criterio de justicia independiente del contrato, alguna manera en que pueda evaluarse la justicia objetiva de un intercambio. Tales obligaciones por lo tanto no son contractuales en el sentido estricto de que el contrato crea la obligación, sino en el sentido epistémico o heurístico limitado de que el contrato ayuda a identificar o aclarar una obligación que ya existía (Atiyah 1979: 143-6). Una consecuencia de esta característica de las obligaciones basadas en el beneficio es que la celebración del contrato no es esencial para la existencia de la obligación. Al menos en principio, pueden existir maneras de identificar tales obligaciones sin recurrir al contrato.

Aquellas obligaciones que surgen del ideal de autonomía, sin embargo, no presumen ninguna calidad de justicia intrín-

seca a ciertos resultados que pueda, incluso en principio, identificarse independientemente del proceso que las produce, o antecedentemente. Con este tipo de obligaciones no puede identificarse ningún resultado como justo sin referencia al procedimiento que de hecho se verificó. Lo que es justo no puede conocerse directamente, porque es por definición el producto de un proceso de cierto tipo; no puede conocerse directamente porque debe ser creado, y hasta que es creado no puede conocerse.

Las maneras contrapuestas en que ambos ideales relacionan al contrato con su justificación pueden ilustrarse a partir de la distinción que hace Rawls entre la justicia procesal pura y la justicia procesal perfecta (o imperfecta). En el caso de la justicia procesal pura «no hay un criterio independiente para el resultado correcto, en su lugar existe un procedimiento correcto o imparcial tal que el resultado sea igualmente correcto o imparcial, sea el que fuere, siempre y cuando se haya observado debidamente el procedimiento» (109). En el caso de la justicia procesal perfecta o imperfecta, por otra parte, «existe una norma independiente para decidir el resultado que es justo» y la cuestión reside simplemente en establecer si es posible encontrar un procedimiento que garantice este resultado (108).

Ahora bien, en tanto un contrato plasma el ideal de autonomía, se aproxima al caso de la justicia procesal pura, en el cual el resultado es justo, sea el que fuere, en virtud del contrato que lo produce. Dentro del ideal de reciprocidad, un contrato es un caso de justicia procesal imperfecta, que de hecho intenta aproximarse a una norma de justicia definida independientemente. Como señala Rawls: «Un rasgo característico de la justicia procesal pura es que el procedimiento para determinar el resultado justo tiene que ser efectivamente observado: ya que en estos casos no existe un criterio independiente por referencia al cual se pueda saber que un determinado resultado es justo. [...] Un procedimiento imparcial *hace imparcial* al resultado sólo cuando es efectivamente observado» [Los subrayados son míos] (109).

La respuesta, entonces, a nuestra pregunta preliminar de cómo se justifican los contratos reales parece ser: «De forma incompleta». Como lo atestigua la coherencia no trivial de la «otra cuestión» («¿Pero es *justo* lo que se acordó?»), los contratos reales no son instrumentos morales autosuficientes sino

que presuponen un trasfondo moral preexistente a la luz de la cual pueden calificarse y evaluarse las obligaciones que surgen de ellos. Aunque puede ser justo, en ciertas circunstancias, obligar a una persona por los términos de su acuerdo anterior, no se deriva necesariamente de su acuerdo que los términos en sí sean justos. El sentido común sugiere varias razones por las cuales, en la práctica, los contratos reales pueden arrojar resultados injustos: una o todas las partes pueden ser coaccionados o sufrir otro tipo de desventaja debido a una posición negociadora desfavorable, o ser víctimas de engaño o de error sobre el valor de las cosas objeto del intercambio, o no tener claras sus propias necesidades e intereses. Pero incluso aunque un contrato no tenga resultados injustos (como por ejemplo cuando se compensan los efectos de tales factores), y aunque la imparcialidad del contrato suministre una razón para su cumplimiento, no puede asumirse que lo que lo *hace* justo es el hecho de que ha sido acordado. Los contratos reales son típicamente casos de justicia procesal imperfecta: la justicia procesal pura aparece raramente, o nunca, en el mundo real.

Contratos y argumentos contractualistas

Rawls probablemente no estaría en desacuerdo con esta formulación. A pesar de la base contractualista de su teoría, no supone que el mero hecho de que exista un acuerdo sea la prueba de su imparcialidad, o que los contratos reales sean instrumentos morales autosuficientes que justifican sus propios resultados, ni que las obligaciones contraídas voluntariamente sean inmunes a la crítica desde el punto de vista de principios de justicia preexistentes. La comprensión del sentido en que la teoría de Rawls no se basa en la noción de contrato como un instrumento de justificación es esencial para la comprensión del sentido en que efectivamente lo hace.

El primer tema a destacar al respecto es que el acuerdo en la posición original no da lugar a obligaciones (al menos no directamente) sino a principios de la justicia. Los principios de la justicia incluyen principios de dos tipos: los «principios para instituciones», que se aplican a la estructura básica de la sociedad, y los «principios para individuos», que formulan los de-

beres y obligaciones de las personas con respecto a las instituciones y entre sí. Los primeros definen qué hace que una institución o una práctica social sean justas, mientras que los segundos especifican los términos en que los individuos están obligados a regirse por ellas.

Los principios para individuos especifican dos maneras diferentes en que las personas están obligadas: como una cuestión de deber natural o de obligación. Los deberes naturales son aquellas demandas morales que se aplican a las personas más allá de su consentimiento, tales como el deber de ayudar a quien lo necesita, no ser cruel, ser justo, etc. Tales deberes son «naturales» en el sentido de que no están vinculados con ninguna institución o acuerdo social en particular, sino que se deben a todos los seres humanos en general (138).

Las obligaciones, por el contrario, describen aquellos vínculos morales que constituimos voluntariamente, ya sea por contrato, por promesa o por otra manifestación del consentimiento. Las obligaciones del cargo público al que se accede voluntariamente constituyen un ejemplo. Pero incluso en el caso de las obligaciones, el consentimiento no es suficiente para crear un vínculo. Otra condición es que la institución o práctica sobre la que se acuerda deben ser justas (o casi justas) es decir, de acuerdo con los dos principios de la justicia. Rawls destaca que, más allá de su dimensión voluntaria, nuestras obligaciones reales no nacen nunca únicamente del consentimiento, sino que inevitablemente presuponen un trasfondo moral antecedente, derivado independientemente, a la luz del cual es siempre posible preguntarse si uno *debió* haber consentido o no.

> Las obligaciones [se producen] sólo si son satisfechas ciertas condiciones básicas. El aceptar, o consentir, instituciones claramente injustas no da lugar a obligación alguna. Es una creencia general que las promesas opresivas [son nulas] *ab initio*. Pero, de un modo similar, los acuerdos sociales injustos son, en sí mismos, un tipo de extorsión, o de violencia, y el consentirlos no es obligatorio. (383)
>
> En particular, no es posible tener obligaciones frente a formas de gobierno autocráticas y arbitrarias. En tales casos no existe el trasfondo necesario para que surjan obligaciones consensuales u otros actos, aunque se expresen así. Los vínculos obligatorios presuponen instituciones justas o, al menos, que sean razonablemente justas, dadas las circunstancias. (135-136)

Incluso las promesas no pueden por sí mismas dar lugar a obligaciones. Rawls distingue entre la regla de la promesa y el principio de fidelidad, y argumenta que la obligación de mantener una promesa no es una consecuencia de la promesa, sino de un principio moral antecedente a la promesa, que se deriva de la teoría de la justicia. «Es esencial, como he dicho antes, distinguir entre la regla de prometer y el principio de fidelidad. La regla es simplemente un convenio constitutivo, mientras que el principio de fidelidad es un principio moral, una consecuencia del principio de imparcialidad. [...] La obligación de mantener una promesa es una consecuencia del principio de fidelidad» (386-387). Como práctica constitutiva o convención, la regla de la promesa es análoga a las reglas legales, o a las reglas de un juego; el que sean o no justas es siempre una cuestión que no puede responderse sin recurso a una norma moral independiente de la práctica. «Hay muchas diferencias en las promesas del mismo modo que las hay en la ley de contratos. El que esta práctica [...] sea justa, ha de ser determinado por los principios de la justicia» (386).

Estrictamente hablando, no son las *promesas* las que obligan, sino el principio de fidelidad que nos obliga a cumplir (algunas de) nuestras promesas, y este principio se deriva de la posición original.

> Incluso la regla de hacer promesas no da lugar por sí misma a una obligación moral. Para considerar las obligaciones fiduciarias, hemos de tomar como premisa el principio de la imparcialidad. Así, de acuerdo con la mayoría de las teorías éticas, la justicia como imparcialidad sostiene que los deberes naturales y las obligaciones se producen en virtud de principios éticos. Estos principios son aquellos que serían elegidos en la posición original. (388-389)

Pero si Rawls no toma los contratos o promesas reales como obligatorios, al menos no en sí mismos, ¿en qué sentido es contractualista su teoría? Es importante hacer una distinción entre el papel del consentimiento en la vida real y su papel en la posición original. Aunque el consentimiento es decisivo en la posición original, juega un papel menos central en nuestros deberes y obligaciones reales. A pesar de su derivación contractualista, los deberes naturales se aplican sin referencia a nuestros actos voluntarios, y el consentimiento que requieren

las obligaciones es en todo caso diferente del consentimiento al que se refiere la posición original.

> Así, aún cuando los principios del deber natural se deriven de un punto de vista contractualista, no presuponen un acto de consentimiento, expreso o tácito, más aún, ni siquiera presuponen un acto voluntario para poder aplicarlos. Estos principios para las personas, así como los principios para las instituciones, son los que se reconocerían en la posición original. Estos principios son concebidos como el resultado de un convenio hipotético. Si su formulación muestra que ninguna acción vinculante, sea consensual o de otro tipo, es un presupuesto de su aplicabilidad, entonces se aplican incondicionalmente. La razón por la cual las obligaciones dependen de los actos voluntarios viene dada por la segunda parte del principio de imparcialidad, que enuncia esta condición. Esto nada tiene que ver con la naturaleza contractual de la justicia como imparcialidad. (139)

Los contratos reales surgen en intercambios o acuerdos cuya justificación debe descansar sobre un principio de la justicia; el contrato hipotético surge de principios de la justicia capaces de evaluar tales acuerdos y definir las consecuencias morales de los contratos en general. Como resultado, el papel que las partes deciden asignar a los acuerdos (reales) es diferente del papel que *su* acuerdo (hipotético) juega en la justificación.

Nozick objeta que es en cierta forma inconsistente que una teoría del contrato produzca principios de la justicia que no asignen plena fuerza justificativa a los intercambios voluntarios. Si los contratos son obligatorios, sugiere Nozick, la teoría de Rawls se equivoca al implicar principios que negarían su fuerza en muchos casos, y si los contratos no son obligatorios, la teoría de Rawls se ve debilitada dado que está fundada en el contrato.

> Los argumentos contractuales comprenden el supuesto de que todo lo que surge de un determinado proceso es justo. Sobre la fuerza de este supuesto fundamental descansa la fuerza de un argumento contractual. Seguramente, entonces, ningún argumento contractual debe ser estructurado de tal manera que haga imposible que los principios procesales sean los principios fundamentales de la justicia distributiva por los cuales juzgar las instituciones de una sociedad: ningún argumento contractual debe estructurarse de tal manera que haga imposible que sus resultados sean de la

misma índole que las suposiciones en que se basa. Si los procesos son lo bastante buenos para fundamentar una teoría, lo son también como posibles resultados de una teoría. No es posible que las cosas se den de ambas maneras. (Nozick 1974: 206-7)

Pero esta objeción omite la distinción entre la justicia procesal imperfecta que describe típicamente nuestros acuerdos reales, y la justicia procesal pura que rige, o al menos se supone que rige, en la posición original. O para plantearlo de otra manera, la objeción confunde los contratos con los argumentos contractualistas. Como observa Rawls, los contratos reales no son argumentos sino hechos sociales, cuyas consecuencias morales dependen de alguna teoría moral, contractualista o no. Es esencial alguna distinción de este tipo, entre el hecho de un acuerdo y las bases de su justificación, si deseamos explicar la coherencia de la «otra cuestión» («¿Pero es *justo* lo que se acordó?»), o para que tenga sentido la cuestión relacionada de si uno *debería* haber consentido o no, o hacer posibles los postulados rivales pero correlativos de la autonomía y reciprocidad en los argumentos sobre la justicia.

En esta perspectiva, no existe contradicción en un argumento contractualista que produzca principios que limiten el papel justificador de los contratos. De hecho, aparece un sentido importante en el cual los resultados de un argumento contractualista *no pueden* ser «del mismo tipo» que el supuesto sobre el que se basan. Ya que si, como argumenta Nozick: «Los argumentos contractuales comprenden el supuesto de que todo lo que surge de un *determinado* proceso es justo» [Los subrayados son míos] (Nozick 1974: 206-7), no parece probable suponer que simplemente cualquier acuerdo, al que se llegue bajo cualesquiera condiciones, pueda producir resultados garantizados como justos. Y una vez que las circunstancias del acuerdo se ven como relevantes para su justificación, no puede alegarse que el acuerdo en sí es la fuente de toda justificación. Reconocer la relevancia de las circunstancias es en sí reconocer una sanción moral independiente del acuerdo por la cual se identifican los rasgos moralmente necesarios de la situación.

Tal como demuestra Rawls en su análisis de las promesas, la fuente de *esta* sanción no puede ser otra promesa o acuerdo (tal como una promesa de mantener las promesas), ya que las credenciales de esta promesa antecedente serían igualmente cuestionables. El respaldo de una promesa (o de un contrato)

debe ser algo más que simplemente otra promesa (o contrato). Debe ser una premisa de un tipo diferente. Esta premisa, que en la teoría contractualista «se parece a un acuerdo para mantener acuerdos posteriores, y que, estrictamente hablando, no puede existir» (390), es lo que intenta brindar el acuerdo hipotético de la posición original. Rawls cree que este mecanismo puede aportar una premisa tal que preserve la apelación voluntarista de la teoría del contrato sin volver a caer en la cuestionable regresión que implica un mero acuerdo sobre acuerdos. Antes de evaluar la solución ofrecida por Rawls, puede resultar útil resumir el problema que la justificación intenta resolver, y considerar brevemente dos soluciones alternativas que Rawls rechaza. De esta manera puede ser viable establecer algunas conexiones entre la explicación de Rawls sobre la justificación y ciertos rasgos centrales del proyecto deontológico.

El liberalismo y la prioridad del procedimiento

Para justificar un intercambio o un acuerdo institucional, no es suficiente demostrar que surge de un acuerdo voluntario entre las partes involucradas, al menos por dos tipos diferentes de razones; uno moral, el otro epistemológico. Aunque Rawls no distingue entre estos argumentos en forma explícita, ambos están implícitos en su explicación, y cada uno sirve para reforzar al otro. Podemos llamar al primero el argumento basado en la contingencia y al segundo el argumento basado en el convencionalismo. El primero nos recuerda al argumento basado en la arbitrariedad, formulado en defensa del principio de diferencia contra las concepciones rivales como las meritocráticas y aristocráticas de la igualdad, que no se adaptaban a la «concepción democrática» de Rawls. En el caso de la justificación, comienza con la observación de que, en la práctica, los acuerdos tienen resultados injustos por una serie de razones, como ya se sugirió: una u otra parte puede haber sido coaccionada o haber estado en una posición de desventaja en el momento de la negociación, o haber sido víctima de un error o falsa información sobre el valor de los objetos que se intercambiaban, o haber estado confundida o errada sobre sus propias necesidades e intereses, o en el caso que se involucraran resultados fu-

turos inciertos, puede haber hecho un mal juicio del riesgo, etc. En algunos de estos casos, especialmente aquéllos que involucran directamente coerción o engaño, podemos estar tentados a decir que el intercambio no fue realmente voluntario, o que el «contrato» es inválido, atribuyendo así la injusticia del resultado a un consentimiento inadecuado. Los libertarios y otros que argumentan que los acuerdos voluntarios son totalmente justificados en sí mismos se preocupan por excluir tales casos invocando distinciones entre influencias coercitivas y no coercitivas, tácticas de negociación legítimas e ilegítimas, amenaza y persuasión, etc. (Nozick 1972, Kronman 1980).

Pero Rawls negaría que, en tanto se conserven algunas influencias moralmente arbitrarias, cualquier distinción de este tipo pueda tener éxito en demarcar un rango de acuerdos justificados en sí mismos. No importa cuán estricta sea la definición de los requisitos de un acuerdo voluntario, el hecho de que las diferentes personas estén en diferentes posiciones asegurará que persistan algunas diferencias de poder o de conocimiento, permitiendo que los acuerdos, incluso los «voluntarios», sean influenciados por factores arbitrarios desde el punto de vista moral. «De alguna manera tenemos que anular los efectos de las contingencias específicas que ponen a los hombres en situaciones desiguales y en tentación de explotar las circunstancias naturales y sociales en su propio provecho» (163).

Incluso los acuerdos voluntarios tienen probabilidades de no alcanzar el ideal de autonomía, de acuerdo con el cual las obligaciones incurridas son asumidas por mí en el sentido estricto del «yo» definido como anterior a sus atributos y fines y por lo tanto libre de determinaciones heterónomas. Sólo este sentido del «yo», con la noción de autonomía que admite, excluye completamente las influencias arbitrarias. La sola exclusión de la coerción no alcanza para justificar un contrato, de la misma manera que la sola exclusión, por ejemplo, de los privilegios de clase no alcanza para justificar una meritocracia. En ambos casos, quedan demasiadas variables sujetas a contingencias arbitrarias desde un punto de vista moral. Una vez que nos afectan los obstáculos más conspicuos a la autonomía individual, es probable que decidamos rechazar las influencias heterónomas siempre que aparezcan.[14]

14. Compárese con Rawls en el tema del principio de la diferencia (1971, 96-97).

Más allá de la dificultad moral con respecto a la noción de que los contratos se justifican por sí mismos yace una dificultad epistemológica. Ésta se refiere a la clasificación de los contratos como «convenciones constitutivas», (la frase es de Rawls), y recuerda el problema que surgiera en conexión con el punto arquimédico, el problema de distinguir una norma de evaluación del objeto de tal evaluación. En el caso de los contratos, la distinción paralela se hace entre el principio moral, por una parte, y la regla, práctica o convención por otra. A pesar de la relevancia normativa de, por ejemplo, la práctica de realizar promesas, las reglas de un juego, o la ley de los contratos, las prácticas, reglas y leyes como éstas no pueden justificar nada por sí mismas, sino que deben depender de algún principio independiente de ellas para justificar sus consecuencias morales. «La doctrina contractual sostiene que de la mera existencia de instituciones no se derivan exigencias morales» (388). Dada su categoría como convención constitutiva, un intercambio no puede justificarse demostrando que fue objeto de un acuerdo voluntario, de la misma manera en que una ley no puede justificarse demostrando que fue aprobada siguiendo el procedimiento adecuado. El hecho de que una transacción sea acordada o de que una ley haya sido aprobada puede ser suficiente, dado el trasfondo normativo relevante, para establecer el requisito legal o institucional de obedecerlas, pero «el que estas exigencias estén conectadas con los deberes y las obligaciones morales es otra cuestión» (389). Todavía necesitamos saber si las partes deberían haber dado su consentimiento o si los legisladores deberían haber votado como lo hicieron.

Si los contratos reales deben presuponer un principio antecedente para justificar sus resultados, surge naturalmente la cuestión de cómo deben derivarse tales principios. Podemos sentirnos tentados a buscar este principio en un acuerdo anterior, más general, que establezca los términos que definen un contrato en particular como justo. Pero esta solución se ve rápidamente socavada una vez que comprendemos que los problemas de las contingencias y los convencionalismos simplemente serían postergados. Una vez que se reconocen estos problemas, no existe ninguna razón para pensar que los contratos de segundo orden pueden justificarse a sí mismos de una manera diferente a como lo hacen los contratos particulares. Un contrato no puede ser sancionado por un acuerdo previo para mantener los contratos, de la misma manera que una

ley no puede justificarse con otra ley sobre la legislación. En ambos casos, la «otra cuestión» no se disuelve sino que se pospone. Una convención acerca de convenciones no constituye un principio moral, sólo otro hecho social.

Si ningún contrato real, sin importar cuán general sea, puede justificar los contratos, debido a su implicación con las prácticas y convenciones de una sociedad en particular, la alternativa parecería ser recurrir a un principio de justicia que fuera de alguna manera anterior a las prácticas y convenciones particulares. Ésta, de hecho, es la solución que buscaban los teóricos contractualistas tradicionales, que respaldaban el contrato social con una apelación al Derecho Natural (Barker 1948: x-xi). Así, para Locke, es «la ley de Dios y de la Naturaleza» la que sanciona el compacto original y establece los límites a los poderes del Estado que surgen de éste.

> Así, el derecho natural se afirma como una regla eterna para todos los hombres, legisladores tanto como los demás. Las reglas que establecen para las acciones de los otros hombres deben conformarse a la ley natural, es decir a la voluntad de Dios, de quien constituyen una declaración (al igual que sus propias acciones, y las de otros hombres). Siendo la ley natural fundamental la preservación de la humanidad, ninguna aprobación humana puede ser válida en su contra. (Locke 1690: 90)

Pero esta solución tradicional no está disponible para Rawls al menos por dos razones. La más obvia es que la confianza en «la ley de Dios y de la Naturaleza» involucra un compromiso teológico y metafísico más substancial del que Rawls está preparado a asumir. Un supuesto tan controvertido entraría en colisión con su determinación de favorecer «condiciones ampliamente aceptadas y preferiblemente débiles» (36), y de «asegurar que los principios de la justicia no dependan de imposiciones más fuertes. [...] Se trata de hacer la menor cantidad posible de suposiciones en la base de la teoría» (155).

A los fines de Rawls, otra dificultad que plantea la idea de basar la justicia en una premisa de derecho natural como la que invoca Locke consiste en que se contradice con el supuesto central del proyecto deontológico. Como hemos visto, una aspiración central del liberalismo deontológico es derivar un grupo de principios regulativos que no presupongan concepción particular del bien alguna, ni dependan de ninguna teo-

ría particular de la motivación humana. Las perspectivas de que las concepciones del bien son diversas, de que no existe un objetivo humano único y dominante, de que el hombre es un ser cuyos fines son elegidos y no dados, y de que una sociedad bien ordenada es por lo tanto una sociedad en la que las personas son libres de perseguir sus distintos objetivos, cualesquiera sean estos, en términos justos aparecen vinculadas a este objetivo. «La libertad para adoptar una concepción del bien está limitada únicamente por principios deducidos de una doctrina que no impone restricciones previas sobre estos conceptos» (290). Para enfatizar la concepción voluntarista de la facultad de agencia humana que subyace a estos principios, Rawls asume que las partes en la posición original no están limitadas por lazos morales anteriores, y «se consideran a sí mismas como seres que pueden elegir y eligen sus últimos fines (siempre varios en número)» (622).

Al apelar al derecho natural como una premisa del compacto original, Locke funda la justicia en ciertos postulados sobre los fines y motivaciones humanas que una ética deontológica sería reticente a admitir. Cuando asume que los hombres se agrupan en sociedad con otros «a fin de unirse para la mutua preservación de sus vidas, libertades y bienes, que denominaré con el nombre general de propiedades», atribuye a los seres humanos un cierto fin dominante y transforma este fin en una premisa de los principios que surgen como resultado. «El gran y principal fin, por lo tanto, detrás de la unión de los hombres en comunidades, y de ponerse bajo un gobierno, es la preservación de su propiedad: con respecto a ella en el estado de naturaleza muchas cosas no existen» (1690: 90).

Pero para Rawls, el fundar principios de la justicia en fines o deseos dados por la naturaleza, ya sean la búsqueda de la felicidad o la preservación de la propiedad o de la vida misma, es ubicar lo bueno antes que lo justo, negando la pluralidad esencial de los fines humanos, y postular en su lugar un único fin dominante; equivale a basar la justicia sobre ciertas contingencias naturales, y revertir la prioridad del «yo» y sus fines al concebir al hombre como un sujeto de fines dados anticipadamente, y no como un sujeto de fines voluntarios que él ha escogido.

Una vez que se admite como premisa del compacto original un fin dominante dado por la naturaleza, las partes ya no pueden concebirse «a sí mismas como seres que pueden elegir y

eligen sus últimos fines», sino limitadas por anticipado. La dimensión voluntarista de su empresa se desdibuja, y los términos del contrato ya no son materia de elección sino que se determinan por anticipado, están dados por los requisitos del derecho natural que constituyen su presupuesto. (Pitkin 1965: 990-9).

El hecho de que Kant sume una objeción similar a las perspectivas contractualistas tradicionales refuerza la idea de que un liberal deontológico rechazara la solución de Locke sobre esta base. Aunque la crítica de Kant está dirigida contra Hobbes, sus objeciones parecerían igualmente aplicables con respecto a Locke.

Kant distingue entre aquellos acuerdos sociales entre seres humanos diseñados para el logro de un fin común que pueden compartir fortuitamente, y aquéllos dedicados a un fin en sí mismo que todos deberían compartir, es decir, el principio de lo justo. Sólo el segundo tipo de contrato constituye un estado civilizado, en el cual la libertad de cada uno se ajusta para armonizarse con la libertad de los demás bajo términos regidos por lo justo. Pero un acuerdo así no puede basarse en ninguna perspectiva particular de la naturaleza humana, ni estar motivada por fines humanos meramente contingentes.

> El concepto completo de un derecho externo se deriva enteramente del concepto de *libertad* en las relaciones externas mutuas entre los seres humanos, y no tiene nada que ver con el fin que todos los hombres tienen por naturaleza (es decir, el objetivo de la búsqueda de la felicidad) o con los medios reconocidos de alcanzar este fin. Así, este último fin no debe interferir de ninguna manera, como determinante, con las leyes que regulan el derecho externo. (Kant 1793: 73)

Para Kant, el principio que funciona como justificación del contrato original no es «el fin principal de preservar la propiedad», ni el de buscar la felicidad, sino el deber en sí mismo que es «la condición formal más alta de todos los demás deberes externos, el derecho del hombre bajo las leyes públicas coercitivas según el cual cada uno puede recibir lo que le corresponde y está seguro frente al ataque de los demás» (Kant 1793: 73). Y este deber no proviene de la naturaleza sino que «es el requisito de la razón pura, que legisla a priori, más allá de cualquier fin empírico (los cuales pueden resumirse bajo el tí-

tulo general de felicidad)». Dado que los seres humanos tienen perspectivas diferentes sobre qué es el fin empírico de la felicidad y en qué consiste, la felicidad no puede someter su voluntad a cualquier ley externa que la armonice con la libertad de todos. El estado civilizado puede solamente establecerse de acuerdo con «principios puramente racionales del derecho externo del hombre» dados a priori. Debe basarse en principios a priori, ya que ni la naturaleza ni la experiencia pueden brindar el conocimiento de lo que es justo. (Kant 1793: 73-4, 86). El resultado es un liberalismo que se aparta significativamente del de Locke, y contiene los puntos esenciales de la ética deontológica que retoma Rawls. En palabras de Kant:

> No nos ocupamos aquí de la felicidad que el sujeto puede esperar obtener de las instituciones o de la administración del bien común, sino primordialmente de los derechos que de esta manera se asegurarían para todos. *Y éste es el principio más alto a partir del cual deben comenzar todas las máximas relativas al bien común, y que no puede ser calificado por ningún otro principio.* Ningún principio de legislación de validez general puede estar basado en la felicidad, ya que tanto las circunstancias del momento como las ilusiones tan conflictivas y diversas sobre lo que es la felicidad (y nadie puede prescribir a los demás cómo deben alcanzarla) hacen imposible cualquier principio firme, de manera que la felicidad por sí sola no puede nunca ser un principio adecuado para la legislación. [Los subrayados son míos.] (Kant 1793: 80)

Mientras que Locke respalda el contrato original con la ley de Dios y de la Naturaleza, Kant lo respalda con un principio de lo justo dado no por la naturaleza sino por la razón pura. De las dos soluciones, la de Kant es la más cercana a la concepción de Rawls en cuanto evita derivar lo justo de lo bueno, y por lo tanto preserva los supuestos deontológicos. Pero como ya hemos visto, Rawls se resiste a la solución de Kant en tanto parece depender de supuestos metafísicos que él encuentra objetables. Duda de la metafísica idealista por la cual funciona la razón pura, y se preocupa por lo que parece ser la arbitrariedad de una derivación apriorística de la ley moral de Kant. De esta manera, en lugar de adoptar la solución de Kant al problema de la justificación directamente, Rawls intenta reformular las enseñanzas deontológicas de Kant, «la estructura subyacente de la doctrina de Kant se separa de sus bases metafísicas, de manera que pueda verse más claramente y se

presente libre de objeciones» (301). Retornamos sobre la misión de la posición original en su intento de suministrar un punto arquimédico; descubrir una forma media entre el convencionalismo y la arbitrariedad, buscar una norma de evaluación que no esté ni comprometida por su implicación en el mundo ni disociada y por lo tanto descalificada por el distanciamiento.

Con respecto a la teoría contractualista, el desafío que plantea el punto arquimédico toma una forma más determinada. Claramente la justificación involucra algún tipo de relación entre los contratos y los principios. Los contratos reales presuponen principios de la justicia que se derivan a su vez de un contrato original hipotético. ¿Pero cómo funciona la justificación en *ese caso*? ¿Se requiere el recurso a otro estrato más de principios antecedentes? ¿O el contrato en esa etapa es moralmente autosuficiente, y completamente justificado en sí mismo? Por momentos la búsqueda de la sanción última parece una danza infinitamente elusiva entre el procedimiento y el principio, que retroceden por turno uno detrás del otro. Dados los supuestos de la teoría contractualista, ninguno parece ofrecer un punto de apoyo sólido sobre el cual fundar al otro. Si las partes en el contrato original *eligen* los principios de la justicia, ¿cómo determinar que hayan escogido *imparcialmente*? Y si eligen a la luz de principios *dados* anticipadamente, ¿en qué sentido puede decirse que hayan sido *elegidos*? La cuestión de la justificación se transforma entonces en una cuestión de prioridad; ¿qué es anterior (en realidad, anterior en última instancia), el contrato o el principio?

En el caso de Kant, no resulta claro si el principio de lo justo es el producto del acuerdo original o su premisa, y su confianza en «la razón pura, que legisla a priori» (1973: 73), parece sugerir esto último. Ésta es en todo caso la cuestión sobre la cual Rawls intenta reformular la posición kantiana, para aseverar la prioridad del contrato y enfatizar así la conexión entre la justicia y la teoría de la elección racional. «El mérito de la terminología contractual es que transmite la idea de que se pueden concebir los principios de la justicia como principios que serían escogidos por personas racionales, y de que las concepciones de la justicia se pueden explicar y justificar de esa manera» (33-34).

Pero al fundar la justicia sobre un contrato original, Rawls trata de expresar lo que toma como la visión kantiana central,

«la idea de que los principios morales son objeto de la elección racional» (288). Describe la posición original como «una interpretación *procesal* de la concepción kantiana de la autonomía y del imperativo categórico [...] un *proceso* natural, traducción de la concepción kantiana del reino de los fines y de las nociones de autonomía y del imperativo categórico» [Los subrayados son míos] (293-301).

¿Por qué una traducción procesal? ¿Por qué es necesario corregir a Kant a fin de dar a estos principios una derivación explícitamente procesal o contractual? La respuesta de Rawls debe ser que al plantear la ley moral como el resultado de un determinado proceso de elección natural, por más hipotético que sea, es posible establecer su postulado sobre la experiencia humana de una forma que de otra manera no resultaría patente. «Estas nociones ya no son puramente trascendentales y faltas de conexión con la conducta humana, ya que la concepción procesal de la posición original nos permite establecer estos lazos» (293).

La prioridad del procedimiento en la concepción de Rawls de la justificación, recuerda las prioridades paralelas de lo justo sobre lo bueno, y del «yo» sobre sus fines. De esta manera vincula la concepción de la justificación con la teoría de la persona que aparece implicada en la justicia como imparcialidad, y sugiere la importancia de la teoría contractualista para el proyecto deontológico en general. De la misma manera en que el «yo» es anterior a los fines que afirma, el contrato es anterior a los principios que genera. Por supuesto, no cualquier contrato es anterior a los principios de la justicia; como hemos visto, los contratos reales no pueden justificarse precisamente porque están típicamente situados en las prácticas y convenciones que la justicia debe evaluar. De forma similar, las personas reales, habitualmente concebidas como «llenas [...] de rasgos particulares», no son estrictamente anteriores con respecto a sus fines, sino que están rodeadas y condicionadas por los valores, intereses y deseos de entre los cuales el «yo» «soberano», en tanto sujeto de la posesión, tomará sus propósitos. Para afirmar la prioridad del yo cuya agencia soberana se asegura, era necesario identificar un yo «esencialmente descarnado», concebido como un puro sujeto de la posesión, distinto de sus objetivos y atributos contingentes, siempre situado detrás de ellos.

En el caso del contrato, la prioridad del procedimiento depende de distinguir un caso especial de justicia procesal pura,

una especie de versión purificada y presituada del procedimiento ordinario en el cual no existe un criterio de imparcialidad independiente. Sólo el «yo» purificado tiene la garantía de ser un agente soberano, y sólo la justicia procesal pura tiene la garantía de producir resultados justos. Es ésta la noción de procedimiento que invoca Rawls en la posición original.

> La intención de la posición original es establecer un procedimiento equitativo según el cual cualesquiera que sean los principios convenidos, éstos sean justos. El objetivo es utilizar la noción de la justicia procesal pura como base de la teoría. (163)

> La justicia procesal pura se da cuando no hay un criterio independiente para el resultado correcto: en su lugar existe un procedimiento correcto o imparcial tal, que el resultado es igualmente correcto o imparcial, sea el que fuere, siempre y cuando se haya observado debidamente el procedimiento. (109)

Otros dos paralelismos vinculan la concepción de Rawls de la justificación con su teoría de la persona. Uno destaca el papel que juega la elección en la ética deontológica: el otro enfatiza el supuesto de la pluralidad. En nuestra argumentación sobre el «yo», consideramos dos explicaciones de la facultad de agencia por las cuales el «yo» podía llegar a sus fines: una concepción voluntarista que relacionaba al «yo» con sus fines como un sujeto voluntario con objetos de elección, y una cognitiva que relacionaba al «yo» con sus fines como un sujeto capaz de comprensión con objetos de conocimiento. Se consideraba que la prioridad del «yo» sobre sus fines requería la concepción voluntarista.

Una vez que imaginamos a las partes en la posición original buscando principios de la justicia, podemos también dar lugar a dos concepciones posibles de la justificación: una voluntarista, en la cual las partes llegan a principios a través de un acto de elección o de acuerdo, y una cognitiva, en la cual las partes llegan a los principios a través de un acto de descubrimiento o percepción colectiva. Sucede lo mismo con la justificación que con el caso de la agencia: para que el contrato sea anterior al principio, las partes deben elegir los principios de la justicia, no descubrirlos. Tanto la prioridad del «yo» como la prioridad del procedimiento exigen nociones voluntaristas de la agencia y de la justificación respectivamente. Para que el «yo» sea anterior, sus fines deben ser elegidos y no dados; para que el con-

trato sea anterior, los principios de la justicia deben ser producto de un acuerdo y no objetos de conocimiento.

Además del énfasis en la elección, el supuesto de la pluralidad es común tanto a la teoría de la persona como a la concepción de la justificación de la perspectiva contractualista. Al igual que «la pluralidad de personas distintas con sistemas de fines separados» (46) es esencial para la noción del sujeto de Rawls, la pluralidad antecedente de las partes en la posición original es esencial para la noción de su acuerdo hipotético, ya que, para que un contrato sea tal, es necesaria una pluralidad de personas. Yo no puedo hacer un acuerdo o contrato conmigo mismo, excepto en un sentido metafórico en el cual parte de la metáfora implica referirnos al «yo» como si fueran dos personas y no una, una pluralidad del «yo» dentro de un solo ser humano. Como hace notar Rawls, los principios de la justicia «se aplican a las relaciones entre varias personas o grupos. La palabra "contrato" sugiere [...] esta pluralidad» (34).

Una vez reconstruido el problema de la justificación tal como lo enfoca Rawls, y después de considerar la forma de solución que emprende, nos resta por demostrar si esta solución es eficaz para aportar los fundamentos que requiere la ética deontológica. Así volvemos finalmente al contrato hipotético en la posición original, a fin de observar qué es lo que pasa exactamente en este caso, y cómo se justifica, si es que lo hace. A riesgo de reiterar nociones que ya nos son conocidos, debemos explorar el texto con cierto detalle para que pueda quedar clara la fenomenología del acuerdo original.

¿Qué sucede realmente tras el velo de la ignorancia?

Lo que se produce en la posición original es en primer lugar una elección, o más precisamente, la elección conjunta de un acuerdo entre partes. Lo que las partes acuerdan son los principios de la justicia. A diferencia de la mayoría de los contratos reales, que no pueden justificarse a sí mismos, el contrato hipotético que las partes acuerdan sí se justifica; los principios que se eligen son justos en virtud de su elección. Como sugeri-

ría la concepción voluntarista de la justificación, los principios de la justicia son el producto de una elección.

> La idea directriz es que los principios de la justicia para la estructura básica de la sociedad son el objeto del *acuerdo* original. [Los subrayados son míos.] (28)

> Así pues, hemos de imaginarnos que aquellos que se entregan a la cooperación social *eligen, en un acto conjunto*, los principios que han de asignar los derechos y deberes básicos y determinar la división de los beneficios sociales. Los hombres habrán de *decidir* anticipadamente cómo regularán las pretensiones de unos y otros, y cuáles serán los principios fundamentales de su sociedad. [Los subrayados son míos.] (28)

> Así como cada persona tiene que *decidir* mediante la reflexión racional lo que constituye su bien, esto es, el sistema de fines que para él es racional perseguir, *del mismo modo un grupo de personas tiene que decidir* de una vez y para siempre lo que para ellas significará justo o injusto. La *elección* que los hombres racionales harían en esta situación hipotética de igual libertad, suponiendo por ahora que este problema de *elección* tiene una solución, determina los principios de la justicia. [Los subrayados son míos.] (28-29)

> Dado que todos están situados de manera semejante y que ninguno es capaz de delinear principios que favorezcan su condición particular, los principios de la justicia serán el resultado de *un acuerdo o un convenio* justo [...] [Los subrayados son míos.] (29)

> La justicia como imparcialidad comienza, como he dicho, con una de las elecciones más generales que las personas pueden hacer en común, esto es, con la *elección* de los primeros principios de una concepción de la justicia [...] [Los subrayados son míos.] (30)

> Los principios de la justicia son aquellos que serían *escogidos* en la posición original. Son el producto de una cierta situación de *elección*. [Los subrayados son míos.] (61)

> En la justicia como imparcialidad los principios de la justicia no son pensados como evidentes por sí mismos, sino que *tienen su justificación en el hecho de que habrán de ser escogidos* [...] [Los subrayados son míos.] (62)

> En una doctrina contractualista los hechos morales están determinados por los principios que deberían ser escogidos en la po-

sición original. [...] *[E]s a las personas* en la posición original a quienes toca *escoger* estos principios [...] [Los subrayados son míos.] (65)

La justicia como imparcialidad difiere de las teorías contractualistas tradicionales en que «el contenido del acuerdo apropiado no es ingresar en una sociedad dada o adoptar una forma dada de gobierno, sino aceptar ciertos principios morales» (33). El resultado del acuerdo no consiste en un grupo de obligaciones que se aplican a los individuos, al menos no directamente, sino en principios de la justicia que se aplican a la estructura básica de la sociedad. De todas maneras el aspecto voluntarista de la justificación se corresponde en cierto sentido con la noción de la sociedad como un acuerdo voluntario. Rawls afirma que vivir en una sociedad regida por los principios de la justicia derivados de una concepción voluntarista de la justificación es, en efecto, lo mejor después de vivir en una sociedad que realmente hayamos elegido.

> Por supuesto que ninguna sociedad puede ser un esquema de cooperación en el cual los hombres ingresen voluntariamente en un sentido literal; cada persona se encuentra desde su nacimiento en una posición determinada de alguna sociedad determinada, y la naturaleza de esta posición afecta materialmente a sus perspectivas de vida. Aún así, una sociedad que satisfaga los principios de la justicia como imparcialidad se acerca en lo posible a un *esquema voluntario*, ya que cumple con los principios que *consentirían* personas libres e iguales bajo condiciones que son imparciales. En este sentido, sus miembros son autónomos y las obligaciones que reconocen son *autoimpuestas*. [Los subrayados son míos.] (30-31)

Tal como lo sugiere nuestra reconstrucción, la naturaleza voluntarista de la perspectiva del contrato de Rawls está cargada con la pluralidad esencial de los sujetos humanos y la necesidad de resolver las demandas en conflicto. Sin la pluralidad, los contratos no serían ni posibles ni necesarios, y en realidad tampoco los principios de la justicia. «Los principios de la justicia se ocupan de las pretensiones conflictivas producto de las ventajas obtenidas por la cooperación social; se aplican a las relaciones entre varias personas o grupos. La palabra "contrato" sugiere tanto esta pluralidad, como la condición de que la división correcta de ventajas tie-

ne que hacerse conforme a principios aceptables para todas las partes» (34).

Como ya hemos visto, la justicia como imparcialidad se diferencia del utilitarismo por su énfasis en la pluralidad y particularidad de los individuos, y esta diferencia está representada en el papel que juega el contrato en la justificación.

> Mientras que el utilitarista hace extensivo a la sociedad el principio de elección de cada hombre, la justicia como imparcialidad, *por ser una visión contractual*, supone que los principios de elección social y por lo tanto los principios de la justicia, son ellos mismos objeto de un *acuerdo* original. [Los subrayados son míos.] (47)

> En todo caso, desde el punto de vista de la *teoría contractualista* no se puede llegar al principio de elección social extendiendo simplemente el principio de la prudencia racional al sistema de deseos construidos por el espectador imparcial. Hacerlo así es no considerar seriamente *la pluralidad y la particularidad de los individuos*, ni reconocer como base de la justicia aquello que los hombres *consintieran*. [Los subrayados son míos.] (48)

Al basar los principios de la justicia en un acuerdo entre partes, Rawls enfatiza dos características que el contrato hipotético comparte con el real, v.g. la elección y la pluralidad. Pero ya hemos visto que los ingredientes de la elección y de la pluralidad no son suficientes para conformar una justicia; los contratos reales, que incluyen ambos, no se justifican a sí mismos. Esto se debe a los problemas que hemos descrito como la contingencia y el convencionalismo. Los contratos reales a menudo tienen resultados injustos debido a diversas contingencias (coercitivas y no coercitivas) que se vinculan con las diferencias inevitables de poder y conocimientos que existen entre personas en diferente situación. Pero en la posición original, tales contingencias se resuelven. Debido al velo de la ignorancia y otras condiciones de igualdad, todas están situadas en forma similar, y por lo tanto ninguna puede aprovecharse, ni siquiera involuntariamente, de una posición negociadora más favorable.

La posición original también está diseñada para superar el problema del convencionalismo. Aunque los contratos reales están inevitablemente insertos en el contexto de las prácticas y convenciones de una sociedad en particular, el acuerdo en la posición original no está implicado de la misma manera. No es

un contrato real, sino sólo uno hipotético. Ya que se imagina qué ocurre antes de que entren en juego los principios de la justicia, puede considerarse como «presituado» en el sentido relevante, un *status quo* antecedente a la llegada de la justicia de forma tal que no se cuenta con principios morales previos por medio de los cuales pudieran impugnarse sus resultados. De esta forma puede realizar el ideal de justicia procesal pura. (Irónicamente, aunque la naturaleza hipotética del acuerdo original parecía en un primer momento debilitar su fuerza justificativa, aparece ahora como una ventaja positiva, incluso indispensable. Si bien Rawls enfatiza que «no es necesario que algo parecido [al acuerdo original] haya existido nunca [...]» (145), puede darse el caso de que un acuerdo como éste no *pudiera* tener lugar nunca y de todos modos superar el problema del convencionalismo.)

> Dado que *todos están situados de manera semejante* y que ninguno es capaz de delinear principios que favorezcan su condición particular, los principios de la justicia serán el resultado de un acuerdo o de un convenio *justo* [...] [Los subrayados son míos.] (29)

> Podría decirse que la posición original es el *status quo inicial apropiado* y que, en consecuencia los acuerdos fundamentales logrados en ellas son *justos*. Esto explica lo apropiado del nombre «justicia como imparcialidad»: trasmite la idea de que los principios de la justicia se acuerdan en *una situación inicial que es justa*. [Los subrayados son míos] (29-30)

> Es un estado de cosas en el cual las partes están igualmente representadas como personas morales y *el resultado [no] está condicionado por contingencias arbitrarias* o por el balance relativo de las fuerzas sociales. Así, la justicia como imparcialidad es capaz de usar la idea de la justicia procesal pura desde el comienzo. [Los subrayados son míos] (145)

Al imponer el velo de la ignorancia es posible «anular los efectos de las contingencias específicas que ponen a los hombres en situaciones desiguales y en la tentación de explotar las circunstancias naturales y sociales en su propio provecho». (163).

> Si se permite un conocimiento de las particularidades, entonces el resultado se ve prejuiciado por *contingencias arbitrarias*. Tal y como se observó, el dar a cada uno según el poder de su ame-

naza, no es un principio de justicia. Si la posición original ha de producir acuerdos que sean justos, las partes tendrán que estar *equitativamente situadas* y tratadas equitativamente en tanto que personas morales. La arbitrariedad del mundo tiene que ser corregida mediante el ajuste de las circunstancias de la situación contractual inicial. [Los subrayados son míos.] (168)

Una vez que las partes de un acuerdo se asumen como situadas en forma similar en todos los aspectos relevantes, desaparecen las diferencias de poder y conocimientos, y así se erradican las posibles fuentes de injusticia. Ya que nadie puede elegir basándose en los atributos dados de manera contingente, se cumple el ideal de autonomía, implícito pero imperfecto en los contratos reales. El ideal de reciprocidad se realiza de forma natural, y la vulnerabilidad del contrato frente a la «otra cuestión» (pero, ¿es justo?) se ve eliminada. «Ahora bien, el velo de la ignorancia priva a las personas en la posición original del conocimiento que les capacitaría para elegir principios heterónomos. Los grupos llegan a su elección conjunta como personas libres y racionales, sabiendo únicamente que aquellas circunstancias dan lugar a la necesidad de los principios de la justicia» (288).

Una vez que la «otra cuestión» de la imparcialidad pierde su fuerza moral independiente, debido al hecho de que las partes están situadas de manera tal que no es concebible que se dé ninguna injusticia, todo acuerdo alcanzado se transforma en un caso de justicia procesal pura; su resultado es justo, «sea cual fuere», en virtud solamente de este acuerdo. Bajo tales condiciones, un contrato deja de ser una convención constitutiva y se transforma en un instrumento de justificación.

El objetivo es caracterizar esta situación de manera que los principios que resultaran escogidos, *cualesquiera que fueran*, fuesen aceptables desde un punto de vista moral. La posición original se define de tal modo que sea un *status quo* en el cual *cualquier acuerdo que se obtenga sea equitativo.* [Los subrayados son míos.] (145)

La intención de la posición original es establecer un procedimiento equitativo según el cual *cualesquiera que sean los principios convenidos, éstos sean justos.* El objetivo es utilizar la noción de la justicia procesal pura como base de la teoría. [Los subrayados son míos.] (163)

Pero en este punto surge una ambigüedad crucial, ya que no resulta claro qué es lo que significa exactamente «utilizar la noción de la justicia procesal pura como base de la teoría». Rawls afirma que *una vez* que la situación es caracterizada adecuadamente, los principios elegidos, *cualesquiera que sean,* son aceptables desde un punto de vista moral; *una vez que* la posición original es definida adecuadamente, *cualquier acuerdo que se obtenga* en ella es justo; *una vez que* se establece un procedimiento justo, *cualesquiera que sean los principios convenidos* serán justos.

Lo que no resulta claro es cómo de generosas deben ser estas provisiones para quienes eligen. En una lectura, los términos parecen realmente generosos, una personificación de las provisiones voluntaristas que se sugirió anteriormente. Una vez que las partes se encuentran en una situación justa, todo vale; el rango de su elección es ilimitado. Los resultados de sus deliberaciones serán moralmente aceptables, «cualesquiera que sean». Sin importar cuáles principios se eligen, éstos se considerarán justos.

Pero existe otra lectura posible de su situación, menos expansiva, que asigna un rango considerablemente menor a esta empresa. Para esta interpretación, decir que los principios escogidos serán justos «cualesquiera que sean» significa simplemente que, dada su situación, está garantizado que las partes escogerán los principios *correctos*. Aunque puede ser cierto que, estrictamente hablando, pueden elegir cualquier principio que deseen, su situación está diseñada de forma tal que está garantizado que «deseen» escoger sólo ciertos principios. En esta perspectiva, «cualquier acuerdo que se obtenga» en la posición original es justo, no porque el procedimiento consagre cualquier resultado como justo, sino porque la situación garantiza un resultado determinado. No obstante, si los principios convenidos son justos porque sólo los principios justos pueden ser convenidos, el aspecto voluntarista de la empresa no es tan amplio como parecería a primera vista. La distinción entre la justicia procesal pura y la justicia procesal perfecta se diluye, y se vuelve menos claro si el procedimiento «traduce su imparcialidad al resultado», o si la imparcialidad del procedimiento se da debido al hecho de que conduce necesariamente al resultado correcto.

Rawls confirma la lectura menos voluntarista cuando afirma que «la aceptación de estos principios no se postula como

una ley psicológica o como una probabilidad. Mi ideal es, en todo caso, mostrar que su aceptación es la única elección [*sic*] coherente con la descripción completa de la posición original. El argumento intenta eventualmente ser estrictamente deductivo» (147). La noción de que una descripción completa de la posición original determina una única «elección» que las partes no pueden sino reconocer parece introducir después de todo un elemento cognitivo a la justificación, y cuestionar la prioridad que tiene el procedimiento sobre el principio que parece requerir la perspectiva contractualista (y el proyecto deontológico en general). Una consecuencia más inmediata de esta lectura es que complica nuestra explicación de lo que sucede en la posición original. Rawls mantiene que lo que sucede detrás del velo de la ignorancia es que una pluralidad de personas llega a un acuerdo unánime sobre una concepción particular de la justicia. Vale la pena examinar esta descripción más de cerca.

> Para empezar, está claro que, puesto que las diferencias entre las partes son desconocidas y puesto que todas son igualmente racionales y se hallan en la misma situación, *todas serán susceptibles de ser convencidas por los mismos argumentos*. Por tanto, podemos contemplar el acuerdo en la posición original desde el punto de vista de una persona seleccionada al azar. *Si cualquiera, después de reflexionar debidamente, prefiere una concepción de justicia a otra, entonces todos lo harán pudiéndose obtener un acuerdo general*. [Los subrayados son míos.] (166)

Rawls sugiere que, para hacer las circunstancias más vívidas, podemos imaginar que las partes se comunican entre sí a través de un árbitro, quien transmite las alternativas propuestas, informa a las partes cuando tienen que llegar a un acuerdo, etc. «Sin embargo, de hecho, tal árbitro es superfluo y se supone que las deliberaciones de los participantes han de ser semejantes» (166).

> Así surge, consecuentemente, el hecho, muy importante, de que las partes no tienen base para regatear, en el sentido usual del término. Nadie conoce su situación en la sociedad ni sus dotes naturales y por lo tanto nadie está en posición de diseñar principios que le sean ventajosos. (166)

> El velo de la ignorancia hace posible la elección unánime de una determinada concepción de la justicia. Sin estos límites pues-

tos al conocimiento, el problema de las negociaciones en la posición original sería infinitamente complicado. (167)

Dado que las partes están «equitativamente situadas», está garantizado que razonarán en forma similar, y no tienen base alguna para negociar «en el sentido usual del término». Esto parecería implicar que tienen una base de negociación en algún otro sentido («los principios de la justicia serán el resultado de un acuerdo o un convenio justo» 29), pero resulta difícil imaginar cuál podría ser este sentido. La negociación en *cualquier* sentido requiere cierta diferencia en los intereses o preferencias, o en el poder o conocimientos de los negociadores, pero en la posición original no hay ninguna. Bajo tales condiciones, es difícil imaginar cómo podría darse negociación alguna, en cualquier sentido.

Si no es posible que se dé negociación alguna, surge entonces la cuestión de si es posible que se produzca el debate. Rawls sugiere que pueden proponerse varias alternativas antes de que se llegue a un acuerdo final. Pero si se asume que las partes razonan de la misma manera y son convencidas por los mismos argumentos, parece poco probable que una cierta idea se le pueda ocurrir a uno pero no a otro. La discusión, tanto como la negociación, presupone alguna diferencia en las percepciones o intereses, o en los conocimientos o preocupaciones de los participantes; sin embargo, en la posición original no existen estas diferencias. Debemos por lo tanto asumir que las «deliberaciones» de las partes tienen lugar en silencio y producen una sola concepción que es acordada en forma unánime.

Sin embargo esto hace que la explicación del acuerdo en la posición original resulte aún más extraña, ya que si no existe una base para la negociación o el debate, resulta dudoso también que pueda haber una base para el acuerdo, incluso para un acuerdo unánime. Rawls declara: «Si cualquiera, después de reflexionar debidamente, prefiere una concepción de justicia a otra, entonces todos lo harán *y* podrá obtenerse un acuerdo general» [Los subrayados son míos.] (166). Pero, ¿por qué «y»? ¿Qué es lo que *agrega* el acuerdo una vez que se ha realizado el descubrimiento? Supongamos que todos, después de reflexionar, descubren que preferirían una determinada concepción de la justicia, y supongamos además que todos los demás supieran que prefieren la misma. ¿El acuerdo sobre esta concepción sería *el próximo paso*? ¿Qué significaría que pri-

mero lleguen a este descubrimiento, y *luego* realicen un acuerdo al respecto? Aunque pudiéramos imaginar lo que significa en tales circunstancias llegar «luego» a un acuerdo, ¿qué *agregaría* el acuerdo al descubrimiento de que todos prefieren la misma concepción? ¿Estaría justificada la concepción, después de que se llegara de alguna manera a un acuerdo, si no quedara justificada en el momento en que se dieran cuenta de que todos prefieren la concepción, sino que ya lo estaba antes de que «lleguen a un acuerdo»?

En este punto es importante trazar una distinción entre dos sentidos diferentes de «acuerdo». La primera implica el acuerdo con una persona (o personas) con respecto a una proposición, la segunda el acuerdo con una proposición. El primer tipo de acuerdo es una forma de «elección conjunta», y requiere una pluralidad de personas. (Una sola no alcanza, excepto en el sentido metafórico en el que yo hago un acuerdo conmigo mismo). Es este tipo de acuerdo el que se da típicamente en la celebración de un contrato, cuando parte del acuerdo implica el formar una intención. Aunque podemos hablar de que dos personas acuerdan *con* un contrato, lo que queremos decir es que dos personas se ponen de acuerdo *entre* ambas para cumplir ciertos términos. El acuerdo y los términos, tomados en conjunto, constituyen el contrato. Dado que el acuerdo en este sentido requiere un acto intencional, o un ejercicio de la voluntad, podemos describirlo como un acuerdo en el sentido voluntarista.

El segundo tipo de acuerdo, el estar de acuerdo con una proposición, no requiere más que una sola persona, y no implica un ejercicio de la voluntad. En este sentido del acuerdo, estar de acuerdo con una proposición implica reconocer su validez, y no requiere ni que otros estén involucrados ni que la validez de la proposición se tome como una cuestión de elección. Puede ser suficiente con que yo la *considere* como válida, como cuando estoy de acuerdo (o acepto, o reconozco) la proposición de que 2 + 2 = 4. Estar de acuerdo en este sentido es comprender algo que ya existía. Aunque puedo decir que he «decidido» que la respuesta a este difícil problema matemático es «*x*», no se trata de una decisión que efectivamente *decida*, sino nada más de si estoy en lo correcto o no. Dado que el acuerdo en este segundo sentido es una cuestión de conocimiento más que de voluntad, podemos describirlo como un acuerdo en el sentido cognitivo.

Una vez que tenemos esta distinción en mente, la explicación de Rawls sobre el acuerdo original aparece en una nueva perspectiva. Los pasajes que parecerían describir un acuerdo en el sentido voluntarista pueden ahora considerarse como admitiendo también una interpretación cognitiva. Aunque al principio Rawls asume que: «La elección [...] *determina* los principios de la justicia» [Los subrayados son míos] (28-29), en otros momentos se expresa como si las partes simplemente tuvieran que *reconocer* principios que ya existen.

> El contenido del acuerdo apropiado no es ingresar en una sociedad dada o adoptar una forma dada de gobierno, sino *aceptar* ciertos principios morales. [Los subrayados son míos.] (33)

> Argumentaré que los dos principios serían *reconocidos* ... [en la posición original]. [Los subrayados son míos.] (143)

> Son los principios que las personas libres y racionales interesadas en promover sus propios intereses *aceptarían* en una posición inicial de igualdad [...] [Los subrayados son míos.] (28)

> [Los miembros de una sociedad bien ordenada] podrían considerar que sus arreglos satisfacen las estipulaciones que hubiesen reconocido en una situación inicial que incorporara restricciones ampliamente aceptadas y razonables para elegir los principios. [Los subrayados son míos.] (30)

> Así los hombres exhiben su libertad, su independencia de las contingencias de la naturaleza y de la sociedad, actuando por medios que *reconocerían* en la posición original. [Los subrayados son míos.] (292)

Irónicamente, la interpretación kantiana de la justicia como imparcialidad destaca el viraje desde la interpretación voluntarista hacia la cognitiva. Aunque se conserva cierta referencia a la elección, las partes se describen más como sujetos que *perciben* el mundo de una cierta manera que como agentes voluntarios.

> Sugiero que concibamos la posición original como un punto de vista a través del cual seres noumenales *contemplan* el mundo. Los grupos compuestos de seres noumenales tienen completa libertad para elegir los principios que deseen; pero también tienen un deseo de expresar su naturaleza de miembros racionales del

reino de lo inteligible donde tienen esta libertad de elección, es decir, como seres que pueden *contemplar el mundo de este modo*, y expresar esta *perspectiva* en su vida como miembros de la sociedad. [Los subrayados son míos.] (292)

Tanto en la teoría de la persona de Rawls como en su explicación de la justificación, los supuestos de la elección y de la pluralidad se han destacado juntos como rasgos centrales de la concepción. A medida que la interpretación voluntarista de la posición original cede ante la cognitiva, se cuestiona también el supuesto de pluralidad. Rawls habla siempre de las *partes* en la posición original y en la interpretación kantiana habla incluso de varios *yoes* noumenales. Pero dado que el velo de la ignorancia tiene el efecto de privar a las partes, en cuanto partes en la posición original, de toda característica distintiva, se vuelve difícil percibir en qué podría consistir su pluralidad.

Rawls reconoce esta condición en parte cuando afirma que «todos están situados de manera semejante» (29), y que de esta manera se garantiza un acuerdo unánime. Pero una vez que se excluyen *todas* las características individualizadoras, las partes no están meramente situadas de manera *semejante* (como las personas del mundo real con circunstancias de vida similares y cierta superposición de intereses), sino de manera *idéntica*. Y como ya hemos visto, la propia teoría de la persona de Rawls reconoce que ningún par de sujetos puede situarse de manera idéntica y seguir considerándose aún como personas diferentes. La noción de que lo que se encuentra tras el velo de la ignorancia no son personas diferentes sino un único sujeto explicaría por qué no puede tener lugar una negociación ni un debate en ese lugar. Explicaría a su vez por qué no puede existir un contrato o acuerdo en el sentido voluntarista, ya que los contratos, como las discusiones, exigen una pluralidad de personas: cuando cae el velo de la ignorancia, la pluralidad se disuelve.

Al principio de su obra, y nuevamente al final, Rawls se pregunta por qué, si la posición original es meramente hipotética, resulta de interés, moral o de otro tipo, para nosotros. Su respuesta en cada caso es que «las condiciones incorporadas a la descripción de esta situación son unas condiciones que nosotros, realmente, aceptamos. O, si no las aceptamos, podemos persuadirnos de hacerlo, mediante las consideraciones fi-

losóficas» (38, también 587). Las consideraciones filosóficas con las cuales nos persuadiría tienen su raíz en la tradición contractualista. La sociedad bien ordenada que recomienda «se acerca en lo posible a un esquema voluntario» (30). Pero lo que comienza como una ética de la elección y del consentimiento, culmina, aunque de manera no deliberada, en una ética del conocimiento y de la autocomprensión. En el pasaje final del libro, el lenguaje de la elección y de la voluntad se ve desplazado por el lenguaje de la percepción y del conocimiento, a medida que la imagen voluntarista de Kant cede lugar ante la imagen cognitiva de Spinoza.

> Una vez que hemos *captado* esta concepción, podemos, en cualquier momento, *mirar* al mundo social desde el *punto de vista* requerido. [...] Así, *observar* nuestro lugar en la sociedad desde la *perspectiva* de esta situación es *observarlo sub specie aeternitatis*: es *contemplar* la situación humana, no sólo desde todos los *puntos de vista* sociales, sino también desde todos los *puntos de vista* temporales. La *perspectiva de la eternidad* no es una *perspectiva* desde un cierto lugar, más allá del mundo, ni el punto de vista de un ser trascendente; más bien, es una cierta *forma de pensamiento* y de sentimiento que las personas racionales pueden adoptar en el mundo. [...] La pureza de corazón, si pudiera alcanzarse, consistiría en *ver claramente* y en actuar con indulgencia y autodominio desde esta posición. [Los subrayados son míos.] (648-649)

El secreto de la posición original —y la clave de su fuerza como justificación— no reside en lo que se hace en ella sino en lo que se aprehende. Lo que importa no es lo que se elige sino lo que se percibe, no lo que se decide sino lo que se descubre. Lo que sucede en la posición original no es después de todo un contrato, sino la conquista de la autoconciencia de un ser intersubjetivo.

IV

La justicia y el bien

Nos hemos propuesto evaluar el postulado de Rawls sobre la primacía de la justicia, y hemos descubierto que exigía una cierta concepción del sujeto moral. Luego tratamos de examinar su concepción a la luz de su teoría moral como un todo para verificar su coherencia con esta teoría y su plausibilidad en general. Esperábamos que de esta manera evaluaríamos primero la teoría de la persona de Rawls y finalmente el postulado de la primacía de la justicia que ésta debe respaldar.

Hasta este momento hemos considerado la teoría del sujeto de Rawls especialmente en relación con su teoría de la justicia, con su concepción de lo justo. Pero como señala él mismo, una teoría moral general debe brindar cierta explicación del bien tanto como de lo justo, y el tercio final de esta obra está dedicado a su búsqueda. Realmente, la primacía de la justicia es en sí un postulado no sólo sobre la justicia sino también sobre la relación de la justicia con las virtudes que se engloban bajo el concepto del bien. Por lo tanto antes de que podamos evaluar este último postulado, debemos considerar la teoría del sujeto de Rawls en relación con esta teoría del bien.

La unidad del yo

Podemos comenzar por recordar los puntos principales de la correspondencia entre la teoría moral y su teoría del sujeto.

Mientras que la moralidad de lo justo concierne a los límites del «yo» y trata de aquello que nos distingue, la moralidad de lo bueno corresponde a la unidad de las personas y trata de aquello que nos une. En una ética deontológica, en la que lo justo es anterior a lo bueno, esto significa que lo que nos separa es, en un sentido importante, anterior a lo que nos une: anterior tanto epistemológica como moralmente. Somos primero individuos distintos, y *luego* formamos relaciones y nos involucramos en acuerdos cooperativos con otros; de ahí la prioridad de la pluralidad sobre la unidad. Somos en primer lugar sujetos despojados de la posesión, y *luego* escogemos los fines que poseeremos; así, la prioridad del «yo» sobre los fines.

Éstos son en resumen los postulados vinculados de la teoría moral y la antropología filosófica sobre la que se ha apoyado el liberalismo deontológico. Al considerar la teoría del sujeto de Rawls desde el punto de vista de lo justo, nos hemos concentrado en la distinción del «yo» y la constitución de sus restricciones. Al considerar la teoría del sujeto ahora desde el punto de vista del bien, nuestro eje estará puesto en la unidad del «yo» y la cuestión de cómo pueden negociarse o atravesarse sus restricciones. Esta cuestión consiste en dos partes. La primera se refiere a la teoría de la comunidad de Rawls y su concepción de cómo las personas anteriormente individualizadas se reúnen para conformar una unión social. La segunda se ocupa de su teoría de la agencia y su explicación de cómo los sujetos de la posesión anticipadamente limitados llegan a adquirir sus propósitos y fines.

Hemos tratado dos rasgos centrales del «yo» —su diferenciación y su unidad— como si cada uno fuera en cierto sentido autosuficiente, como si cada uno pudiera ser descripto en forma independiente del otro. Sin embargo es difícil en la práctica observar la distinción entre estos dos rasgos del «yo» sin destacar también su vinculación interna. Incluso al habernos concentrado en los límites del «yo» y en la concepción de lo justo, ya hemos tenido ocasión de buscar un principio de unidad y de anticipar una teoría de la comunidad que pueda aportarlo. Incluso hemos visto desde el punto de vista de lo justo cómo ceden los límites del «yo» que establece Rawls.

En tanto se ha considerado que el principio de diferencia requiere de un sujeto de la posesión más amplio, los principios de la justicia han rebasado los límites del sujeto anticipada-

mente individualizado, por así decirlo, y se han basado anticipadamente en una forma de unidad que Rawls reserva a la provincia de lo bueno. Mientras que el autor fijaría las identidades de las personas independientemente de su condición común, y definiría lo justo sin referencia a (una teoría general de) lo bueno, la noción de que el principio de diferencia se basa en una teoría de la comunidad «desde el principio» negaría estas prioridades de maneras sustanciales. Mientras que Rawls considera que la teoría del bien no es un requisito previo sino un complemento de la justicia, diseñada para demostrar su estabilidad y su tendencia a generar sus propios apoyos, la noción más amplia del sujeto de la posesión que está implícita en el principio de diferencia requeriría ciertas teorías de la comunidad y de la agencia en la base de la justicia y no en su perímetro.

Necesitamos por lo tanto evaluar la teoría del bien de Rawls, y en especial sus concepciones de la comunidad y de la agencia, no sólo con respecto a su plausibilidad general, sino también en cuanto a su capacidad de ofrecer cierta explicación que requiere la teoría de la justicia para ser completa. Intentaré demostrar que la concepción de Rawls no logra ninguno de estos aspectos, y por razones similares. Pero antes de concentrarnos directamente en su teoría del bien, puede resultarnos útil considerar una ilustración concreta de qué está exactamente en juego para la justicia en la teoría de la comunidad, y qué es lo que no funciona en la posición liberal cuando trata de prescindir de ella. Con este fin me propongo considerar un argumento de Ronald Dworkin a favor de la acción afirmativa, o tratamiento favorecido de las minorías en el ingreso a las universidades. Aunque el argumento de Dworkin no es idéntico al que propondría Rawls, tiene mucho en común con la perspectiva general de Rawls sobre el mérito y la naturaleza del sujeto moral, y sirve para destacar los supuestos deontológicos de los que nos ocupamos.

La defensa de la acción afirmativa

Dworkin defiende las políticas de ingreso basadas en la acción afirmativa, en el caso de las universidades como las de

medicina o derecho, sobre la base de que son un medio eficaz, o al menos potencialmente eficaz, de lograr un objetivo social deseable, el de aumentar la presencia de personas negras y otras minorías en tales profesiones socialmente estratégicas, y de esa manera eventualmente «reducir el grado en que la sociedad estadounidense es sobre todo una sociedad pendiente del tema racial» (1977b: 11). Su argumento básico se centra en la utilidad social. La acción afirmativa se justifica, no porque aquellos que son favorecidos tengan derecho a una ventaja, ya sea en compensación por la discriminación del pasado o por cualquier otra razón, sino simplemente porque «el ayudarlos es hoy una forma eficaz de atacar un problema nacional» (1977b: 12).

Pero Dworkin, como Rawls, cree que ninguna política social puede ser justificada, por mucho que contribuya al bienestar general, si viola los derechos individuales. Por lo tanto, tiene en cuenta el argumento de que la acción afirmativa viola los derechos de aquellas personas blancas que pone en desventaja y que en algunos casos excluye. Llega a una conclusión negativa: la idea de que el tratamiento preferente «presenta un conflicto entre un objetivo social deseable e importantes derechos individuales, surge de una confusión intelectual [...]» (1977b: 12).

Una versión del argumento que considera Dworkin es el postulado de que el tener en cuenta la raza es injusto porque se concentra en una cualidad más allá del control de la persona. Dworkin responde que esto no es privativo de la raza como criterio, sino que se aplica igualmente a la mayoría de los criterios que se utilizan típicamente en la admisión a facultades y universidades, incluyendo la inteligencia. Aunque es verdad que una persona no elige su raza:

> [...] también es verdad que aquellos que obtienen bajas calificaciones en los exámenes de aptitud o de ingreso no eligen sus niveles de inteligencia. Ni tampoco eligen no tener las cualidades que exigen a aquellos a los que se niega el ingreso debido a que sobrepasan la edad, o porque vienen de una parte del país que no esté suficientemente representada en la facultad, o debido a que no pueden jugar bien al baloncesto [...] (1977b: 15).

La raza puede parecer un factor diferente debido a que las exclusiones basadas en ella han expresado históricamente pre-

juicio o desprecio por la raza excluida como tal. Pero las personas blancas excluidas debido a la acción afirmativa no lo son por desprecio sino simplemente por el mismo tipo de cálculo instrumental que justifica los criterios que nos resultan más conocidas. Aunque es verdad que una persona blanca con una calificación marginal en los exámenes hubiera sido aceptada si fuera negra, «también es verdad, y en exactamente el mismo sentido, que hubiera sido aceptado si hubiera sido más inteligente, o hubiera causado una mejor impresión en la entrevista [...] La raza no es, en *su* caso, una cuestión diferente de estos otros factores, igualmente más allá de su control» (1977b: 15).

Otra versión del argumento que considera Dworkin es el postulado de que al admitir personas negras con calificaciones más bajas que las que logran algunas personas blancas que son excluidas, la acción afirmativa viola el derecho de los solicitantes a ser evaluados sobre la base del mérito. La respuesta de Dworkin es que lo que se considera mérito no puede determinarse en abstracto sino que depende de las cualidades consideradas relevantes para el propósito social que cumple la institución. En el caso de las facultades de medicina y de derecho, la inteligencia medida por pruebas estandarizadas puede estar entre las características relevantes, pero de ninguna manera es el único criterio adecuado, como lo demuestra una práctica bien establecida de los comités de admisión. Otros atributos de la persona y de su currículum se tienen típicamente en cuenta al evaluar la probable habilidad del solicitante para realizar la función necesaria, y en el caso de que el ser negro fuera relevante para el propósito social del que se trata, el serlo debe contar también como mérito.

> No hay una combinación de habilidades, destrezas y características de la personalidad que constituya el «mérito» en abstracto: si la destreza manual se considera «mérito» en el caso de un futuro cirujano, esto se debe a que tal capacidad le permitirá servir al público mejor que a ninguna otra causa. Si el ser negro, admitiendo que es un hecho lamentable, le permitirá a otro médico realizar mejor un trabajo diferente dentro de la medicina, entonces el serlo se considerará «mérito» por el mismo criterio. (1977b: 13)

Dworkin reconoce que algunas personas pueden considerar el argumento que considera a la raza como una forma de mé-

rito peligroso, pero «sólo porque confunden su conclusión —que la raza negra puede ser una característica socialmente útil en circunstancias particulares— con la idea muy diferente, e innoble, de que una raza puede ser intrínsecamente más valiosa que otra» (1977b: 12).

La idea que se encuentra implícita en gran parte de la obra de Dworkin es que nadie puede alegar imparcialmente que sus derechos son violados por los programas de acción afirmativa, porque ninguna persona, blanca o negra, *merece* ir a la facultad de medicina o de abogacía por principio: nadie tiene un derecho anterior a ser admitido. Seguramente, aquellos que cumplen de la forma más completa las condiciones establecidas para la admisión tienen *un derecho* a ser admitidos, y sería incorrecto excluirlos. Pero no puede decirse que ellos o ninguna otra persona *merezcan* ser admitidos, al menos por dos razones. Primero, el que tengan las características relevantes no es en la mayoría de los casos resultado de su propio esfuerzo; su inteligencia innata, su entorno familiar, sus oportunidades sociales y culturales, etc., son en la mayoría factores más allá de su control, una cuestión de buena suerte. Y en todo caso, en primer lugar nadie tiene derecho a que las facultades de medicina o de derecho se comprometan a recompensar ningún tipo de cualidades. Lo que se considera una cualidad para una tarea en particular depende de las aptitudes que la tarea requiera, y de nada más. Los beneficios asociados con las profesiones no son por lo tanto recompensas por logros superiores sino incentivos para atraer las cualidades relevantes. Por lo tanto no puede existir un derecho anterior a ser evaluado de acuerdo con ningún conjunto de criterios en particular.

A partir de esto resulta claro que los argumentos de Dworkin coinciden con la teoría de Rawls en varios aspectos. La noción de que los criterios tradicionales de admisión, al igual que la raza, no son resultado del esfuerzo de los solicitantes, nos recuerda al argumento de Rawls de que las ventajas de los privilegiados son arbitrarias desde un punto de vista moral. El argumento de Dworkin de que no existe el «mérito» en abstracto, sin referencia a los propósitos que las instituciones pueden definir y perseguir, es paralelo al de Rawls que se opone a la meritocracia en cuanto a que los conceptos del mérito, de la virtud y del valor moral no tienen un rango moral preinstitucional o anterior, y por lo tanto no pueden constituir un punto de partida independiente desde el cual puedan criticar-

se instituciones de otra manera justas. Y la implicación general del argumento de Dworkin de que ninguna persona, blanca o negra, merece asistir a la facultad de medicina o de derecho, y de que nadie tiene un derecho anterior a ser admitido, se corresponde con la distinción que traza Rawls entre el mérito moral y las expectativas legítimas.

Las posiciones de Rawls y de Dworkin son similares también en un sentido más general. Ambas son teorías basadas en derechos, definidas en oposición explícita a las concepciones utilitaristas, y buscan defender ciertas demandas individuales contra el cálculo del interés social. Pero a pesar de sus aspiraciones individualistas, ambas se basan en una teoría del sujeto que tiene el efecto paradójico de confirmar la fragilidad última, e incluso quizás la inconsistencia, del individuo cuyos derechos intentan sobre todo asegurar. Hemos visto ya cómo la concepción del «yo» de Rawls amenaza en distintos puntos del argumento con disolverse en un sujeto radicalmente incorpóreo, o con derrumbarse como un sujeto radicalmente situado. Como veremos, el argumento de Dworkin a favor de la acción afirmativa ilustra cómo estas perplejidades, identificadas en primer lugar en lo abstracto, tienen consecuencias en la práctica.

En cualquier defensa de la acción afirmativa resulta central la capacidad de distinguir entre la discriminación contra la raza negra y otras minorías, como en las prohibiciones históricas de acceso para las personas de color y las cuotas en contra de los judíos, y la discriminación a favor de las personas negras y otras minorías del tipo que implican los programas de acción afirmativa. Dworkin alega que la justificación del primer tipo de discriminación depende típicamente en parte de la «la idea muy diferente, e innoble, de que una raza puede ser intrínsecamente más valiosa que otra» (1977b: 12), mientras que la justificación de la segunda depende por el contrario de la noción utilitarista de que la sociedad como un todo se beneficiaría al tener profesionales de la medicina y del derecho más ampliamente representativos.

Con respecto a la primera justificación, Rawls rechazaría claramente como Dworkin la idea de que una raza pueda ser intrínsecamente más valiosa que otra. Lo que resulta interesante recordar es por qué, al menos en la teoría del sujeto de Rawls, debe ser incorrecta esta idea innoble. Para Rawls, la falacia del postulado de que las personas blancas son intrín-

secamente más valiosos que las personas negras no reside en que niegue el valor inherente a las personas negras, sino en que atribuya falsamente un valor inherente a las personas blancas, y por lo tanto les atribuya una demanda infundada al mérito. La razón radica en que, para Rawls, el concepto de valor moral, como el concepto del bien, es «inferior a los conceptos de lo correcto y la justicia, y no desempeña ningún papel en la definición substantiva de las porciones distributivas» (352). Las personas no pueden tener un valor intrínseco de la misma manera que no pueden tener un mérito intrínseco, es decir, un valor o mérito que sea suyo anteriormente o independientemente de lo que las instituciones justas les atribuyan. Y como hemos visto, en un sentido estricto, no puede decirse que nadie merezca nada porque no puede decirse que nadie *tenga* nada, o al menos no en el sentido no distanciado y constitutivo de la posesión que es necesario como base del mérito. En la teoría del sujeto de Rawls, ninguna persona o raza puede ser intrínsecamente más valiosa o merecedora que otra, no porque todas sean intrínsecamente valiosas o merecedoras —en forma igualitaria— sino porque ninguna lo es, y por lo tanto todas las demandas deben esperar igualmente el advenimiento de las instituciones justas.

Surgen objeciones al argumento de Dworkin a favor de la acción afirmativa, y a la teoría de la justicia de Rawls en tanto lo respalda, sobre la base meritocrática de que el individuo posee sus atributos en un sentido no problemático y por lo tanto merece los beneficios que se derivan de ellos, y de que parte de lo que significa que una institución o esquema distributivo sean justos es que premian a los individuos previamente merecedores de una recompensa. Pero Rawls y Dworkin presentan poderosos argumentos en contra de estos supuestos, que los defensores de la meritocracia encontrarían difíciles de rebatir. La dificultad del argumento de Dworkin, a mi entender, es otra: se refiere a las posibles visiones alternativas del sujeto que permanecen una vez que se rechaza la concepción meritocrática del individuo. Y esto nos devuelve al problema de los límites del «yo».

Ya hemos considerado las dificultades que se asocian con la noción de una persona esencialmente desposeída, que carezca de rasgos constitutivos, sin valor o mérito intrínseco, y cuyas perspectivas de vida dependan totalmente de los derechos y oportunidades que puedan dispensar las instituciones de la

justicia. Hemos señalado también la ironía de que una persona tan despojada moralmente pueda ser el producto de una ética liberal diseñada para establecer que los derechos de los individuos son inviolables. Pero si bien la negación del mérito individual y la insistencia en los límites entre el «yo» y sus atributos nos conducen en dirección a un sujeto radicalmente no corporeizado, la noción del acervo común plantea una amenaza diferente para la integridad del «yo» en tanto implica que en cierta manera los límites entre el «yo» y los otros deben relajarse. Y es que, a menos que podamos encontrar un principio de individualización diferente al meramente empírico, el peligro consiste en una desviación hacia un sujeto radicalmente situado.

En el argumento de Dworkin a favor de la acción afirmativa, esta perplejidad adopta la siguiente forma: una vez que la admisión o exclusión no pueden ser consideradas de manera plausible como dependientes de una noción abstracta de «mérito» ni de una demanda individual antecedente, la alternativa es asumir que deben prevalecer automáticamente los fines colectivos de la sociedad como un todo. Pero los límites de la sociedad relevante nunca se establecen, su rango como el sujeto de la posesión adecuado nunca se confirma. Una vez que el «yo», en tanto «yo» individual, se ve desposeído, las demandas del individuo se desvanecen para revelar un utilitarismo subyacente que nunca se ve justificado. Y, como Rawls indica anteriormente, el utilitarismo es en cierto sentido la ética de un sujeto no limitado, la ética que no llega a considerar seriamente las distinciones entre las personas.

Para Dworkin, sin embargo, las consideraciones utilitaristas son precisamente las que distinguen a la discriminación legítima que implica la acción afirmativa de la forma no justificable que se basa en el prejuicio y el desprecio. Si no puede decirse que algunos son intrínsecamente más *merecedores* que otros, al menos puede decirse que algunos son más *valiosos* que otros con respecto a los propósitos sociales en cuestión, y la discriminación sobre esta base es justificable. En tanto una política de tratamiento preferente *utiliza* a las personas para lograr fines valiosos en lugar de *juzgar* a las personas como más o menos dignas en sí mismas, resulta permisible. En tanto una exclusión racial esté motivada no por el prejuicio sino por un «cálculo instrumental», un «cálculo racional sobre el uso más beneficioso para la sociedad de los recursos escasos»,

o una idea del tipo «el pertenecer a la raza negra puede ser un rasgo socialmente útil» (1977b: 12), la exclusión resulta coherente con las consideraciones utilitaristas y puede ser justificable. Las expectativas de una persona, a menos que se funden en derechos en el sentido especial que Dworkin atribuye al término, deben siempre ceder frente a «cuestiones sociales más generales», como cuando un pequeño empresario debe desaparecer para que pueda construirse una nueva y mejor autopista (1977b: 15). Aunque su decepción sea comprensible, e incluso merezca nuestra simpatía, los solicitantes rechazados no pueden interponerse en el camino de la profesión médica que la sociedad necesita, de la misma manera que el pequeño empresario no puede anteponerse a la necesidad de la autopista.

Aunque el argumento de Dworkin supone que cuando no están en juego derechos individuales la política social se decide adecuadamente sobre una base utilitarista, nunca explica el porqué de esto. Además de demostrar por qué los argumentos utilitaristas no pueden atacar derechos individuales, su teoría no ofrece una defensa explícita de la ética utilitarista en tanto tal, y poco dice de por qué debería prevalecer el utilitarismo cuando no están implicados derechos individuales. Dworkin puede no sentir la necesidad de justificar sus supuestos utilitaristas subyacentes debido a que en la superficie parecen tener cierta evidente sugestión. Si ningún individuo tiene una demanda antecedente a los beneficios de sus atributos y activos accidentalmente dados, puede parecer natural suponer que la sociedad como un todo sí la tiene. Pero como hemos visto en la argumentación sobre el acervo común y el principio de diferencia, este supuesto no está fundamentado. La arbitrariedad de los atributos de un individuo sólo aporta un argumento en contra de la proposición de que el individuo los posee o que tiene un derecho preferente sobre sus beneficios, pero no a favor de la proposición de que cierta sociedad en particular los posee o tiene un derecho preferente con respecto a ellos. Y a menos que podamos establecer esta segunda proposición, parecería no haber fundamento para favorecer una dispensa utilitarista de tales atributos y activos en lugar de simplemente dejarlos donde están.

Si prescindiéramos de cierta concepción de un sujeto de la posesión más amplio, como la noción de Rawls del acervo común parece también requerir, no existirían razones obvias de

por qué tales atributos deban ponerse al servicio de fines sociales generales en lugar de individuales. Por el contrario; en ausencia de un sujeto de la posesión más amplio, el considerar «mis» atributos y activos como meros instrumentos de un propósito social más amplio es utilizarme como un medio para los fines de otro, y por lo tanto violar un precepto moral central tanto para Rawls como para Kant.

La peculiaridad moral de basar la admisión a las universidades sobre los supuestos que sugieren Rawls y Dworkin, esté o no involucrada la acción afirmativa, puede ilustrarse imaginando las siguientes cartas de rechazo y de aceptación redactadas para ilustrar la base moral de la política que recomiendan:

> Estimado solicitante (no admitido):
>
> Lamentamos tener que informarle de que su solicitud de admisión ha sido rechazada. Rogamos que comprenda que no hay ofensa alguna implícita en nuestra decisión. Su rechazo no indica que lo despreciemos ni siquiera que lo consideremos menos merecedor de ser admitido que aquellos que han sido aceptados.
>
> No es culpa suya que, cuando Ud. se presentó, la sociedad no necesitara las cualidades que Ud. puede ofrecer. Quienes fueron admitidos en su lugar no eran en sí mismos merecedores de una vacante, ni dignos de elogio por los factores que condujeron a su admisión. En última instancia simplemente los estamos utilizando —como a Ud.— como instrumentos de un propósito social más amplio.
>
> Es probable que para Ud. esta noticia sea una decepción en el sentido de que no se cumplirán sus esperanzas de obtener los beneficios que se conceden a aquellos cuyas cualidades coinciden con las necesidades de la sociedad en un momento determinado. Pero este tipo de decepción ocurre siempre que las preferencias de un individuo deben ceder ante las de la sociedad, y no deben exagerarse por el pensamiento de que su rechazo refleja de alguna manera su valor moral intrínseco; le aseguramos que aquellos que fueron admitidos son tan carentes de valor intrínseco como Ud.
>
> Cuente con nuestra simpatía en el sentido de que lamentamos que no haya tenido las cualidades que casualmente la sociedad necesitaba cuando Ud. se postuló. Le deseamos mejor suerte la próxima vez.
>
> Atentamente

Estimado solicitante (admitido):

Nos complace informarle de que su solicitud de admisión ha sido aceptada. Aunque no sea el resultado de su propio esfuerzo, resulta que casualmente Ud. posee los rasgos que la sociedad necesita en este momento, por lo que nos proponemos explotar sus atributos para beneficio de la sociedad admitiéndolo como estudiante de medicina/derecho.

No debe entender o inferir alabanza alguna de esta decisión, ya que el hecho de que posea las cualidades relevantes es arbitrario desde un punto de vista moral. Debemos felicitarlo, no en el sentido de que merezca algún crédito por tener las cualidades que condujeron a su admisión —ya que no es así— sino sólo en el sentido en que se felicita al ganador de la lotería. Tiene suerte de haber llegado con las características adecuadas en el momento adecuado, y si decide aceptar nuestra oferta tendrá en última instancia derecho a los beneficios que se asocian a que lo utilicemos de esta manera. Por todo esto, Ud. puede celebrarlo adecuadamente.

Ud., o más probablemente sus padres, pueden sentirse tentados a celebrarlo también si toman su admisión como una prueba, si no de sus cualidades innatas, al menos del esfuerzo consciente que ha realizado por cultivar sus habilidades y superar los obstáculos para estos logros. Pero el supuesto de que merezca siquiera el carácter superior necesario para su esfuerzo es igualmente problemático, ya que su carácter también depende de diversos tipos de circunstancias afortunadas sobre las cuales Ud. no puede atribuirse crédito alguno. La noción de mérito parece no aplicarse a su caso.

De todas maneras esperamos verlo en el otoño.

Atentamente

Como sugieren estas cartas, la política que defienden Rawls y Dworkin puede resultar problemática incluso para aquellos que no sostienen los supuestos meritocráticos que ellos efectivamente cuestionan. Podemos imaginar, por ejemplo, una respuesta dentro de las siguientes líneas:

Yo no alego que yo, como individuo, posea (en ningún sentido exclusivo) los atributos que se me atribuyen, ni que tenga ninguna demanda moral especial sobre los frutos de su ejercicio. Reconozco que tengo una deuda —con mis padres, mi familia, mi ciudad, mi tribu, mi clase, mi nación, mi cultura, mi período histórico, posiblemente con Dios— en una variedad compleja de maneras con respecto a la constitución de mi identidad, y por lo tanto puedo reclamar muy poco o ningún

crédito (ni tampoco culpa) por haber resultado de la manera en que lo hice. El decidir exactamente quién o qué es el responsable de esto o de determinada parte de mí es una actividad moral difícil, aunque a veces indispensable, que más allá de cierto punto puede volverse imposible de realizar. Pero estoy en todo caso de acuerdo en que no merezco ser admitido para ninguna oportunidad en particular en ningún sentido moral antecedente, primero porque no poseo por derecho propio las cualidades que me harían elegible, y segundo, porque aunque lo hiciera, no tendría derecho a que las normas vigentes recompensen ningún grupo de atributos o cualidades en particular por encima de otros.

A partir de esto es razonable suponer que lo que a primera vista parecen ser «mis» atributos pueden describirse más adecuadamente como un acervo común en cierto sentido: ya que son otros los que me hicieron la persona que soy, y en muchas maneras continúan haciéndolo, parece adecuado considerarlos, en la medida en que puedo identificarlos, como participantes de «mis» logros y beneficiarios comunes de las recompensas que se deriven de ellos. En tanto este sentido de participación en los logros y esfuerzos de (ciertos) otros involucra la autocomprensión reflexiva de los participantes, hemos llegado a considerarnos, sobre el abanico de nuestras diversas actividades, menos como sujetos individualizados con algunas cosas en común, y más como miembros de una subjetividad más amplia (pero aún determinada); menos como «otros» y más como participantes de una identidad común, sea una familia, una comunidad, una clase, un pueblo o una nación.

Una consecuencia de una autocomprensión ampliada como ésta es que cuando «mis» atributos o perspectivas de vida se enrolan al servicio de un esfuerzo común, es probable que experimente esto no tanto como el ser utilizado para los fines de otros sino como una manera de contribuir a los propósitos de la comunidad que considero propia. La justificación de mi sacrificio, si es que puede llamarse sacrificio, no reside en la seguridad abstracta de que otros desconocidos obtendrán un beneficio mayor de lo que yo pierdo, sino en la noción más alentadora de que con mis esfuerzos contribuyo a la realización de una forma de vida de la que estoy orgulloso y con la cual está ligada mi identidad. Aunque por supuesto sigue siendo verdad que yo no podría, como individuo, alegar crédito alguno por poseer las cualidades relevantes al esfuerzo co-

mún, puedo sin embargo estar orgulloso de mi aptitud para contribuir de esta manera; y esta aptitud, quizás incluso más que los beneficios que puedo recibir, será un motivo justo para celebrar.

Esto, por supuesto, no equivale a afirmar que pueda describirse de esta forma una demanda sobre «mis» recursos que provenga de cualquier parte. El alcance de mis lazos comunitarios, por expansivo que sea, no es ilimitado. Incluso un «yo» ampliado, concebido como una comunidad, tiene sus límites, no importa cuán provisionales sean sus contornos. Los límites entre el «yo» y (ciertos) otros, por lo tanto, se relajan en la explicación intersubjetiva, pero no se relajan tan completamente que cedan a un sujeto radicalmente situado. Los límites que permanecen no están dados por las diferencias físicas o corpóreas entre los seres humanos individuales, sino por la capacidad del «yo» de participar a través de la reflexión en la constitución de su identidad, y en tanto lo permitan las circunstancias, llegar a una autocomprensión expansiva.

Otro rasgo de la descripción intersubjetiva del acervo común es que la desposesión de la persona, tal como aparece desde el punto de vista individualista, equivale en última instancia a una pérdida menor de competencias. Aunque el argumento basado en la arbitrariedad priva sistemáticamente al sujeto, en tanto persona individual, de sus atributos y posesiones, dejándolo tan despojado de rasgos empíricamente identificables que se desvanece en la abstracción («La persona ha desaparecido. Sólo quedan los atributos»), la noción de un sujeto de la posesión más amplio adelanta en algo la reconstitución de la persona y la devolución de sus potenciales. Si yo no puedo ser el dueño, al menos puedo ser el guardián de los atributos ubicados «aquí», y lo que es más, el guardián en nombre de la comunidad de la cual me cuento como miembro.

Nada de esto constituye un argumento en contra de la acción afirmativa en sí. Pero sugiere una nueva cuestión moral que Dworkin debe solucionar antes de que su argumento a favor de la acción afirmativa se considere completo, y es la cuestión de cómo establecer el sujeto de la posesión relevante, o cómo identificar a aquellos entre quienes se pueden considerar correctamente como comunes los atributos que accidentalmente se me atribuyen. Para decirlo en otras palabras, el utilitarismo es una ética de la participación. (En este aspecto, nos recuerda al principio de diferencia). En tanto tal debe pre-

suponer cierto lazo o vínculo antecedente entre aquellos cuyas satisfacciones maximizaría y cuyos esfuerzos y expectativas distribuiría en el proceso. De otra manera es simplemente una fórmula para utilizar a algunos como medios para los fines de otros, una fórmula que los liberales deontológicos están comprometidos a rechazar.

Pero la posición de Dworkin con respecto a esta cuestión es por lo menos ambigua. Por momentos se expresa como si no se requiriera explicación alguna de un sujeto de la posesión ampliado, como si fuera suficiente para el éxito de un argumento utilitarista que una expectativa individual se enfrentara a «un propósito social más amplio», en donde la expectativa no está protegida como una cuestión de derecho. En esta interpretación, yo debo compartir «mis» atributos con «la sociedad como un todo», no porque esta sociedad en particular me haya hecho lo que soy y por lo tanto sea responsable de estos atributos y dones de una manera que yo, individualmente, no lo soy, sino basándose en el dudoso supuesto de que la «sociedad» es la beneficiaria residual de todos los atributos que circulan libremente y permanecen, después de que el individuo es desposeído. Esto supone sin discusión que la «sociedad» en cierto sentido indeterminado (¿toda la humanidad?) tiene un derecho anterior sobre todos los atributos que un individuo no tiene. Pero simplemente porque yo, como individuo, no tenga un derecho preferente sobre los activos que residen accidentalmente «aquí», eso no significa que todos en el mundo lo posean colectivamente. Ya que no existe razón para pensar por adelantado que su ubicación en el ámbito de la «sociedad» (o, por decirlo así, en el contexto de la humanidad) sea menos arbitraria desde el punto de vista moral. Y si su arbitrariedad dentro de *mí* hace que no sean elegibles para servir a *mis* fines, no parece haber ninguna razón evidente por la que su arbitrariedad dentro de una sociedad en particular no pueda volverlos inelegibles también para servir a los fines de esa sociedad.

Dworkin se expresa en otros momentos como si después de todo tuviera en mente un determinado sujeto de la posesión, y éste fuera el Estado nación. Afirma, por ejemplo, que «la *sociedad estadounidense* es en la actualidad una sociedad pendiente del tema racial», y que el objetivo de la acción afirmativa es «reducir el grado en que la sociedad *estadounidense* es sobre todo una sociedad pendiente del tema racial». Se dice que estos pro-

gramas brindan «una manera eficaz de atacar un problema *nacional*» [Los subrayados son míos] (1977b: 11-12). Pero si Dworkin tiene la intención de hacer esta afirmación con respecto a la determinación del ingreso a la universidad y las perspectivas de carrera, debería decir bastante más sobre el porqué de todo esto. Y parte de este argumento debería incluir cierta evidencia de la responsabilidad de la nación por haber cultivado las cualidades y atributos que ahora absorbería, su capacidad de abocarse a la autocomprensión reflexiva de sus miembros como la base de su identidad común, y la habilidad de demandar, si no el acuerdo, al menos la aceptación para los fines que surgirían de esta identidad. Necesitaría demostrar, en resumen, que de las diversas comunidades y formas de identidad, la nación es la que tiene un derecho adecuado a definir el propósito común y a destacar los atributos comunes necesarios para su logro, al menos con respecto a la educación universitaria y la elección de ciertas carreras profesionales. Puede darse el caso o no de que la nación estadounidense actual defina una comunidad en el sentido relevante;[15] pero en tanto Dworkin desee invocar a la nación como el sujeto de la posesión relevante, es él el que debe demostrar que esto se verifica.

A pesar de las referencias casuales sobre la nación, tanto Dworkin como Rawls parecen asumir en general que, una vez que se han definido los derechos del individuo, predomina una demanda social no especificada y no se requiere explicación alguna acerca de determinadas comunidades o sujetos de posesión en sentido más amplio. Así, Dworkin se refiere a la necesidad de servir a una «preocupación social más general», y de brindar «lo que la sociedad en general más necesita» (1977b: 15); y Rawls escribe sobre la necesidad de acordar esquemas distributivos diciendo que deben ser de tal manera que propicien «el interés común» (350) y sirvan a «fines sociales prioritarios e independientes» (352).

Podemos resumir las dificultades que plantea este supuesto de la siguiente manera. En primer lugar, no existe nada pa-

15. Se encontrará una exposición iluminadora sobre la nación como comunidad en Beer (1966). Éste distingue entre la centralización del gobierno y el proceso de integración nacional, y señala que las dos tendencias, por más mutuamente dependientes que sean, no necesariamente coinciden. En el caso de la integración nacional, la nación está «más compuesta que una comunidad», y el sentido en el cual sus miembros comparten una vida común se profundiza (1966: 80-82).

recido a «*la* sociedad como un todo» o a «*la* sociedad más general», en abstracto, ni una única comunidad «última» cuya preeminencia se dé sin mayor explicación o sin cuestionarla. Cada uno de nosotros se mueve en una cantidad indefinida de comunidades, algunas de ellas más inclusivas que otras, cada una de las cuales impone demandas diferentes sobre nuestra lealtad, y no hay forma de determinar anticipadamente cuál es *la* sociedad o comunidad cuyos propósitos deberían regir la disposición de cualquier grupo determinado de nuestros atributos y dones.

En segundo lugar, si no existe nada parecido a «*la* sociedad como un todo», en abstracto, parecería poco probable que una sociedad en particular, identificada en forma arbitraria, pudiera tener una demanda más firme sobre un grupo determinado de atributos que el individuo en donde éstos residen accidentalmente, ya que evidentemente su ubicación dentro del ámbito de tal comunidad arbitrariamente identificada no sería menos arbitraria desde un punto de vista moral. En particular no existirían razones obvias de por qué una «preocupación social más general» en sí misma debería prevalecer en todos los casos sobre preocupaciones más locales o particulares meramente en virtud de su generalidad. Resulta interesante notar a este respecto que el utilitarismo en sus versiones más tempranas y teológicas (como las de Tucker y Paley) ofrecía una concepción explícita del sujeto último de la posesión—en este caso Dios— cuyos propósitos predominaban necesariamente sobre las preocupaciones más locales (MacIntyre 1967: 462-6). Pero una vez que el utilitarismo se volvió secular, el sujeto de la posesión relevante deja de ser una cuestión resuelta, y los fundamentos para afirmar la precedencia de una categoría de cuestiones sobre otra aguardan una descripción posterior del sujeto o comunidad relevante y la base de sus demandas.

Finalmente, a menos que sea posible identificar a la comunidad relevante dentro de la cual «mis» atributos se comparten adecuadamente, y establecer sus fundamentos, el argumento de Dworkin a favor de la acción afirmativa y la noción de Rawls del acervo común tienen el efecto, o bien de contradecir el postulado central en Kant y Rawls en contra de la utilización de algunos como medios para los fines de otros, o bien de evadir esta contradicción al relajar completamente los límites entre el «yo» y los otros, cayendo así en un sujeto radicalmente situado.

Habiendo visto una ilustración práctica de cómo la justicia, en la ética deontológica, requiere una noción de comunidad para ser coherente y no simplemente para demostrar su «congruencia y estabilidad», debemos considerar ahora si la concepción de Rawls está en condiciones de proporcionarla. Rawls afirma que «la justicia como imparcialidad ocupa un lugar importante en la valoración de la comunidad [...]» (301). La cuestión radica en determinar si la teoría de la comunidad que ofrece Rawls es capaz de completar los principios de la justicia que se requieren y de explicar también la virtud de la comunidad en general.

Tres concepciones de comunidad

En la evaluación de la teoría de la comunidad de Rawls puede resultarnos útil recordar que el individualismo de su concepción describe al sujeto y no al objeto de las motivaciones. Los intereses perseguidos por las partes en la posición original no son los intereses *en* el «yo» sino los intereses *de* un «yo», y más específicamente los intereses de un «yo» anteriormente individualizado. Al identificar el individualismo de su teoría con el sujeto, en lugar del objeto de deseo, Rawls cree poder evitar basarse en una teoría de las motivaciones humanas en particular, y en especial en el supuesto, común a algunas teorías liberales tradicionales, de que el hombre es por naturaleza egoísta. Derivar una teoría de la justicia sin referencia a ninguna motivación o concepción del bien en particular es esencial para el proyecto deontológico y tiene la consecuencia ulterior, en opinión de Rawls, de permitir una teoría de la comunidad más general de la que es factible en los supuestos individualistas tradicionales. Cuando se deja abierto el contenido de las motivaciones, es posible suponer que los individuos pueden perseguir objetivos sociales o comunitarios al igual que privados, especialmente en una sociedad regida por un esquema de reciprocidad que funciona para afirmar su sentido de autoestima.

> No hay razón alguna por la que una sociedad bien ordenada debiera alentar valores primariamente individualistas si esto significara una forma de vida que conduce a los individuos a ocupar-

se de sus propios intereses y a no mostrar preocupación alguna por los intereses de los demás (aunque respetando sus derechos y libertades). Normalmente es de esperar que la mayoría de las personas pertenezcan a una o más asociaciones, o que tengan al menos algunos fines colectivos en este sentido. (Rawls 1975: 550)

En su argumentación sobre la «idea de la unión social» (sección 79), Rawls distingue dos sentidos para el «bien de la comunidad». El primero se basa en supuestos individualistas convencionales que dan por sentadas las motivaciones egoístas de los agentes. Esta explicación concibe a la comunidad en términos completamente instrumentales y evoca la imagen de una «sociedad privada», en donde los individuos consideran los acuerdos sociales como una carga necesaria y sólo cooperan en aras de la prosecución de sus fines privados. A partir de esta concepción instrumental, Rawls distingue su propia perspectiva de la comunidad en la cual los participantes tienen ciertos «fines últimos compartidos» y consideran el esquema de cooperación como un bien en sí mismo. Sus intereses no son uniformemente antagónicos sino que en algunos casos resultan complementarios o se superponen. Dado que Rawls no asume por anticipado que todos se aboquen únicamente a motivaciones egoístas, no elimina la posibilidad de que algunos puedan tener en cuenta el bienestar de otros e intenten promoverlo. «Por supuesto que no es necesario suponer [...] las personas no hacen sacrificios importantes unas por otras, ya que, por el contrario, los hacen a menudo movidas por el afecto y por vínculos sentimentales. No obstante, tales acciones no son exigidas por motivos de justicia por la estructura básica de la sociedad» (207).

De las dos concepciones de la comunidad que presenta Rawls, ambas son individualistas, aunque la manera en que lo son difiere en cada caso. La explicación instrumental es individualista en el sentido de que los sujetos de la cooperación se suponen gobernados por motivaciones egoístas únicamente, y el bien de la comunidad consiste solamente en las ventajas que los individuos derivan de la cooperación para la prosecución de sus fines egoístas. La explicación de Rawls es individualista en el de sentido que supone la individualización antecedente de los sujetos de la cooperación, cuyas motivaciones reales pueden incluir objetivos fruto de la benevolencia al igual que otros fruto del egoísmo. Como resultado, el bien de la comunidad

para Rawls consiste no sólo en los beneficios directos de la cooperación social sino también en la calidad de las motivaciones y los vínculos sentimentales que esta cooperación puede lograr, resultando mejorados en el proceso. Mientras que en la primera concepción la comunidad es totalmente *externa* a los objetivos e intereses de los individuos que la componen, en la perspectiva de Rawls la comunidad es parcialmente *interna* a los sujetos, en el sentido de que alcanza los sentimientos y emociones de aquellos involucrados en el esquema cooperativo. Por contraste con la concepción instrumental de la comunidad, podemos por lo tanto describir la concepción de Rawls como la concepción sentimental.

Pero ni la explicación instrumental ni la sentimental parecen capaces de generar la teoría fuerte de la comunidad que exigen aparentemente los argumentos de Rawls y de Dworkin; el de Rawls para redimir la noción del acervo común involucrada en el principio de diferencia, el de Dworkin para definir la comunidad de participación relevante para su argumento a favor de la acción afirmativa. Como hemos visto, para estar completos ambos argumentos parecen requerir un sujeto de la posesión más amplio, capaz de plantear una demanda legítima sobre el acervo necesario para sus propósitos sin utilizar a algunos como medios para los fines de otros y sin caer en un sujeto radicalmente situado. Pero ni la concepción instrumental ni la sentimental de la comunidad, que presuponen la individualización anterior del sujeto, pueden ofrecer una manera en la que los límites del sujeto sean replanteados; ninguna de las dos parece capaz de relajar los límites entre el «yo» y los otros sin producir un sujeto radicalmente situado.

Debido a esto, tendríamos que imaginar una concepción de la comunidad que pudiera penetrar más profundamente en el «yo» de lo que permite incluso la concepción sentimental, ya que aunque Rawls admite que el bien de la comunidad puede ser interno en la medida que involucra los objetivos y valores del «yo», no puede serlo en un sentido tan cabal que alcance más allá de las motivaciones al sujeto de éstas. El bien de la comunidad no puede llegar *tan* lejos, ya que el hacerlo sería una violación de la prioridad del «yo» sobre sus fines, la negación de su individualización antecedente, el retroceso de la prioridad de la pluralidad sobre la unidad, y la admisión de que el bien participa en la constitución del «yo», lo cual en la perspectiva de Rawls está reservado al concepto de lo justo.

(«La unidad esencial del «yo» es facilitada ya por la concepción de lo justo», 622).

Una teoría de la comunidad cuyo ámbito incluye al sujeto tanto como al objeto de las motivaciones sería individualista en un sentido distinto tanto al convencional como al de Rawls. Se parecería a la concepción de Rawls en cuanto a que el sentido de la comunidad se manifestaría en los objetivos y valores de los participantes —en la forma de sentimientos fraternales y camaradería, por ejemplo— pero se diferenciaría en que la comunidad describiría no sólo un *sentimiento* sino también una manera de autocomprensión parcialmente constitutiva de la identidad del agente. Dentro de esta perspectiva fuerte, afirmar que los miembros de una sociedad están limitados por un sentido de comunidad no equivale simplemente a declarar que una mayoría de ellos profesa sentimientos comunitarios y persigue objetivos comunitarios, sino que ésos miembros conciben su identidad (el sujeto y no sólo el objeto de sus sentimientos y aspiraciones) como definida en cierta medida por la comunidad de la cual forman parte. Para ellos, la comunidad describe no sólo lo que *tienen* como conciudadanos, sino también lo que *son*; no una relación que eligen (como en la asociación voluntaria) sino un lazo que descubren; no meramente un atributo sino un componente de su identidad. Por oposición a las concepciones instrumental y sentimental de la comunidad, podemos describir esta perspectiva fuerte como la concepción constitutiva.

A pesar de la resistencia de Rawls a la concepción constitutiva de la comunidad y a la teoría del sujeto que requiere, ya hemos visto cómo su lenguaje parece por momentos llevarlo más allá de la concepción sentimental, como si implícitamente reconociera que lo que hemos alegado es verdadero, que su teoría de la justicia depende en última instancia para su coherencia precisamente de la dimensión intersubjetiva que oficialmente rechaza. En la concepción del principio de diferencia se nos dice que la distribución de talentos naturales se describe mejor como un «acervo común», y que en la justicia como imparcialidad los seres humanos acuerdan «compartir el destino común». En la explicación de la unión social, los límites entre las personas empíricas y corpóreas parecen aún más atenuados. Se afirma que los seres humanos tienen «fines últimos compartidos» y participan a través de la comunidad «en la suma total de los valores naturales realizados de los otros».

Por lo tanto se nos conduce a «la noción de la comunidad de la humanidad», cuyos límites podemos imaginar que se extienden incluso a través del tiempo y del espacio, «porque los que se encuentran muy separados por la historia y por las circunstancias pueden, sin embargo, cooperar en la realización de su naturaleza común» (582). «Sólo en la unión social está completo el individuo», ya que es allí donde «cesamos de ser meros fragmentos» (584). Los miembros de la comunidad «participan de sus naturalezas recíprocas», y «el "yo" se realiza en las actividades de muchos» (624).

No obstante, como sugiere la distinción entre las concepciones sentimental y constitutiva de la comunidad, el vocabulario moral de la comunidad en el sentido fuerte no puede ser capturado en todos los casos por una concepción que «en su base teórica es individualista». Así, una «comunidad» no siempre puede traducirse, sin una pérdida de significado, como una «asociación», ni un «vínculo» como una «relación»; ni «compartidos» como «recíprocos», ni «participación» como «cooperación», ni lo que es «común» como lo que es «colectivo». Aunque el argumento de Rawls de la prioridad de la pluralidad sobre la unidad puede normalmente aplicarse al segundo de cada uno de esos pares, no necesariamente se aplica al primero. Mientras que activos «colectivos» implican bienes que una vez se tuvieron en forma separada, y que fueron cedidos a la sociedad como un todo, un acervo «común» no es así necesariamente: no necesita presuponer lógicamente una individualización previa. Y aunque «reciprocidad» implica un principio de intercambio y por lo tanto una pluralidad de agentes, la noción de «compartir» puede sugerir una solidaridad tal que no necesite involucrarse ningún intercambio, como en el compartir una broma, o una aspiración, o un entendimiento. Y mientras que la característica de la «asociación» y la «cooperación» es que presuponen una pluralidad anterior a aquéllos que se reúnen para asociarse o cooperar, la «comunidad» y la «participación» pueden describir una forma de vida en la cual los miembros se encuentren situados conjuntamente «para empezar», y donde esta calidad de estar juntos consiste no tanto en la relación que han establecido como en los vínculos que han encontrado.

De esta manera parecería que la comunidad en el sentido fuerte y constitutivo que requieren tanto Rawls como Dworkin no puede explicarse por una concepción que sea individualis-

ta, ni siquiera en el sentido especial que Rawls da al término, ya que una concepción individualista toma los límites del sujeto como dados previamente, y finalmente fijos. Sin embargo, ambos autores exigen una concepción capaz de demarcar un sujeto de la posesión más amplio, en la cual el sujeto esté potenciado para participar en la constitución de su identidad. A continuación veremos que tal concepción no está disponible en los supuestos deontológicos.

Para que un sujeto participe en la definición de los contornos de su identidad, es necesaria cierta facultad de reflexión. La voluntad por sí sola no es suficiente. Lo que se requiere es una cierta capacidad de conocimiento de sí mismo, una capacidad para lo que hemos llamado agencia en un sentido cognitivo. Esto puede verse recordando las dos concepciones de agencia y posesión que consideramos en nuestra reconstrucción inicial de la teoría del sujeto de Rawls. La primera, que corresponde a su concepción, tomaba los límites del «yo» como dados y relacionaba al «yo» con sus fines por medio de la agencia en su sentido voluntarista, un sujeto de voluntad para los objetos de su elección. Este tipo de agencia dependía de la facultad de la voluntad, ya que es la voluntad la que permite que el «yo» se extienda más allá de sí, que trascienda los límites que están fijos por anticipado, que alcance los fines que poseería y que los mantenga como siempre debe, externos a sí.

La segunda, por el contrario, tomaba los límites del «yo» como abiertos y concebía la identidad del sujeto como el producto y no como la premisa de la agencia. La agencia relevante en este caso no era la voluntarista sino la cognitiva: el «yo» llegaba a sus fines no por elección sino por reflexión, como un sujeto de conocimiento (o indagación) para un objeto de (auto)comprensión. El problema en este caso no era la distancia entre el «yo» y sus fines, sino por el contrario el hecho de que el «yo», al no estar limitado por anticipado, estaba atravesado por propósitos y fines posibles, todos los cuales asaltaban indiscriminadamente su identidad, amenazando siempre con absorberla. El desafío para el agente residía en distinguir los límites o restricciones del «yo», diferenciar al sujeto de su situación, y de esta manera forjar su identidad.

Para el sujeto cuya identidad se constituye a la luz de los fines que ya se encuentran ante él, la agencia consiste más en lograr la autocomprensión que en reunir la voluntad. A diferencia de la capacidad de elección, que permite al «yo» exten-

derse más allá de sí mismo, la capacidad de reflexión le permite al «yo» volver su foco de atención hacia dentro de sí, indagar en su naturaleza constitutiva, investigar sus diversos vínculos y reconocer sus respectivas demandas, distinguir los límites —por momentos expansivos, por momentos reducidos— entre el «yo» y los otros, llegar a una autocompresión menos opaca si bien nunca completamente transparente, una subjetividad menos fluida si bien nunca finalmente fija, y así gradualmente, a lo largo de la vida, participar en la constitución de su identidad.

La capacidad de reflexionar sugerida por la concepción cognitiva parecería precisamente el rasgo que requiere el «sujeto de la posesión más amplio» de Rawls, para no disolverse en un sujeto radicalmente situado, ya que esta capacidad detiene la posibilidad de llegar a los límites del «yo» sin tomarlos como dados por anticipado. De hecho, una vez que se cuestiona la individualización anterior presunta del sujeto, el predicamento del «yo» parecería aproximarse a la exposición descripta en la concepción *cognitiva,* en la cual la mayor amenaza para la facultad de agencia no es la distancia entre el «yo» y sus fines y propósitos, sino el exceso de objetivos aparentemente indispensables que sólo un sobrio examen de sí mismo puede confiar en distinguir.

Sin embargo, en la epistemología moral de Rawls el alcance para la reflexión aparecería seriamente limitado. El conocimiento de sí mismo no parece ser una posibilidad en el sentido relevante, ya que los límites que definiría se toman como dados por anticipado, sin reflexión, de una vez y para siempre, por el principio de la individualización antecedente. Pero una vez que se considera que esos límites se desvanecen, no hay nada que tome su lugar, ya que para un sujeto como el de Rawls la cuestión moral paradigmática no es «¿Quién soy?» —dado que la respuesta a esta pregunta se considera evidente en sí misma— sino «¿Qué fines debo elegir?», y éste es un interrogante que se dirige a la voluntad. El sujeto de Rawls aparece así epistemológicamente empobrecido en lo que respecta al «yo», conceptualmente mal equipado para dedicarse al tipo de reflexión capaz de llevarlo más allá de una atención a sus preferencias y deseos, y para contemplar, y de esta manera volver a describir, al sujeto que los contiene.

Parece claro al menos que la cuestión de la comunidad nos lleva naturalmente a la cuestión de la reflexión, y que para

evaluar el papel de la reflexión en el esquema de Rawls, necesitamos examinar en mayor detalle su teoría de la agencia, su concepción de la manera en que el «yo» llega a sus fines. Hemos visto que para este autor el «yo» es anterior a sus fines por elección, o para decirlo de manera más elaborada, el «yo» está relacionado con sus fines como un sujeto de voluntad con los objetos de elección, y hemos descrito esta capacidad de elegir como *la agencia* en el sentido voluntarista. Pero, ¿qué sucede exactamente en este momento de elección, y qué papel juega la reflexión para llegar a éste, si es que juega alguno?

La facultad de agencia y el papel de la reflexión

Para Rawls, la concepción de la agencia y de los fines se encuentra bajo el concepto del bien. Como lo justo, lo bueno se concibe desde lo voluntarista: se basa en la elección. Así como los principios de lo justo son el producto de una elección colectiva en la posición original, las concepciones del bien son productos de elecciones individuales en el mundo real.

Pero aquí se nos plantea una importante diferencia ya que, aunque tanto lo justo como lo bueno se fundan en la elección, las condiciones especiales (v.g. hipotéticas) bajo las cuales se «elige» lo justo significan que las personas reales no participan en el proceso. Lo que se considera justo o correcto no es algo que tenemos la libertad de elegir, ya que los principios de la justicia están en vigencia desde el momento en que desaparece el velo de la ignorancia, es decir, antes de que pueda comenzar ninguna elección *real*. Los principios de la justicia, al ser derivados en forma antecedente, no están sujetos a nuestra agencia; se aplican nos guste o no.

El caso del bien es diferente. En este caso, cada persona es libre de escoger por sí misma, libre de adoptar cualquier concepción del bien que desee. Personas distintas consideran que son buenas cosas diferentes y, con el único límite de las restricciones de la justicia, cada uno es «libre de proyectar su vida según le plazca» (493). Aunque se asume que existe una única «correcta concepción de la justicia, desde un punto de vista filosófico» (493), a la que todos se deben adherir, no existe una concepción correcta del bien comparable desde un pun-

to de vista filosófico, y por lo tanto cada uno es libre de establecer la suya propia.

Aquí encontramos por lo tanto una nueva ilustración de la prioridad de lo justo sobre lo bueno, tanto en su dimensión moral como en la epistemológica. La prioridad moral consiste en el hecho de que los principios de la justicia limitan las concepciones del bien que los individuos pueden elegir; en caso de que los valores de una persona se opongan a la justicia, es la justicia la que prevalece. Como reconoce Rawls en repetidas ocasiones, los principios de la justicia no son compatibles con todos los planes de vida concebibles, y los planes que no se adecuan deben rechazarse.

> En la justicia como imparcialidad, el concepto de lo justo es prioritario respecto al del bien. En contraste con las teorías teleológicas, algo es bueno sólo cuando se ajusta a las formas de vida compatibles con los principios de lo justo ya existentes. (438)

> Ciertamente, también los planes racionales de vida que determinan qué cosas son buenas para los seres humanos —los valores de la vida humana, por así decirlo— se hallan sometidos, a su vez, a los principios de la justicia. (440)

> Nuestro modo de vida, cualesquiera que sean nuestras circunstancias particulares, debe conformarse siempre a los principios de la justicia a los que se llega independientemente. (496)

> Estos principios reciben, pues, una primacía absoluta, de modo que regulan las instituciones sociales sin discusión, y cada uno construye sus proyectos de acuerdo con ellos. Los proyectos que no se encuentran en esta línea deben ser revisados. (624)

> Los deseos y aspiraciones [de los seres humanos] se restringen desde el comienzo mediante los principios de la justicia que especifican los límites que los sistemas de fines de los hombres tienen que respetar. [...] La prioridad de la justicia se explica, en parte, sosteniendo que los intereses que exigen la violación de la justicia carecen de valor. Al no tener méritos en el primer término, no pueden sobreponerse a sus pretensiones. (50)

La prioridad de lo justo sobre lo bueno brinda un respaldo metaético a la conocida noción liberal de que las preferencias y convicciones de la mayoría, no importa cuán intensas sean, no pueden superar una demanda legítima de los derechos in-

dividuales. En una ética deontológica, las convicciones de la mayoría reflejan meramente concepciones particulares del bien. Como tales, no pueden postular que son «correctas desde un punto de vista filosófico», sino sólo preferidas desde un punto de vista mayoritario, y ninguna mera preferencia puede sobreponerse a los requisitos de la justicia.

> Las profundas convicciones de la mayoría, si son, en realidad, simples preferencias sin fundamentarse en los principios de la justicia previamente establecidos, no tienen suficiente peso. La satisfacción de estos sentimientos carece de todo valor que pueda colocarse en el platillo de la balanza contra las pretensiones de una libertad igual. [...] Frente a estos principios, nada importa la intensidad de los sentimientos ni el hecho de que sean compartidos por la mayoría. (497)

Otra forma de considerar el contraste entre lo justo y lo bueno es recordar que lo bueno, ya sea individual o colectivo, incluye como ingredientes diversas contingencias que son arbitrarias desde un punto de vista moral, mientras que lo justo está libre de tales arbitrariedades. La necesidad que Rawls destaca de regular la distribución de las oportunidades y los beneficios, de forma tal que no resulte arbitraria desde un punto de vista moral, sugiere por lo menos una razón de por qué lo justo debe ser anterior. Una vez que se asegura la precedencia de lo justo, sin embargo, no existe objeción a la prosecución de concepciones del bien teñidas de contingencias y arbitrariedades. Los principios de la justicia domestican eficazmente tales concepciones y las mantienen dentro de los límites en forma segura.

> No hay objeción alguna que hacer al ajuste de los proyectos racionales a estas contingencias, porque los principios de la justicia han sido elegidos ya y determinan el contenido de estos proyectos, los objetivos que fomentan y los medios de que se valen. (496)
>
> Así, los aspectos arbitrarios de los proyectos de vida no afectan a estos principios, ni a la forma en que ha de disponerse la estructura básica. La indeterminación en la noción de racionalidad no se traduce en las legítimas aspiraciones que los hombres pueden imponerse mutuamente. [...] Como las aspiraciones de los hombres no se ven afectadas, la indeterminación es relativamente inocua. (496, 623)

La prioridad de lo justo puede, finalmente, considerarse en términos de su derivación anterior, y la necesidad de algún entorno último «no elegido» como la condición previa de la elección en las concepciones del bien. Si los principios de la justicia fueran en sí mismos materia de elección, ya no estaría asegurada «la libertad de elección que la justicia como imparcialidad asegura a los individuos y grupos dentro de la estructura de unas instituciones justas» (494). Algo debe permanecer fuera de la elección (y de esta manera limitarla), a fin de preservar la elección misma. Ésta es la prioridad epistemológica que la ética deontológica transfiere a una prioridad moral.[16] De esta manera la moralidad de lo justo, que asegura la libertad de elección dentro de sus límites, no puede ser vulnerable *en sí* frente a cualquier elección que la amenazaría o restringiría, ya que la fuerza moral de tal amenaza no alcanzaría en ningún caso a la fuerza moral de la estructura dentro de la cual fue concebida. Tal como lo hemos visto en nuestra exposición sobre la teoría contractualista, la sanción de esta estructura, a diferencia de la sanción de los objetivos y valores que surgen dentro de ella, no es una simple elección, ni siquiera un contrato, sino un acuerdo hipotético concebido bajo condiciones especiales cuya imparcialidad está establecida independientemente.

El aspecto epistemológico de la prioridad de lo justo nos recuerda a la prioridad paralela del «yo» sobre sus fines. En ambos casos se considera que un requisito previo para la elección consiste en un marco de referencia «no elegido», dado anteriormente. Como los principios de la justicia deben estar dados anteriormente (y por lo tanto colocados más allá de la elección) a fin de asegurar la posibilidad de elección en las concepciones del bien, de la misma manera los límites del «yo» deben estar dados antecedentemente (y más allá de la elección) a fin de asegurar la agencia del sujeto, su capacidad de elegir sus fines. Mientras que los límites impuestos por la justicia pueden parecernos una restricción impropia sobre la elección en cuanto a que las personas no pueden participar en

16. Compárese el argumento de Kant en la deducción trascendental de que los conceptos de los objetos en general deben ser subyacentes a todo el conocimiento empírico como su condición a priori. «La validez objetiva de las categorías a priori se halla, por lo tanto, en el hecho de que, en lo que respecta a la forma del pensamiento, sólo a través de ellas la experiencia se vuelve posible» (Kant 1781: 126).

su constitución sino sólo escoger dentro de ellos, de hecho estos límites aseguran la igual libertad de cada uno para elegir sus fines por sí mismo frente a la vaguedad de una opinión pública que un día puede preferir algo distinto. De manera similar, mientras que los límites del «yo» pueden parecernos una restricción impropia sobre la agencia en el sentido de que el «yo» no puede participar en su constitución, de hecho se trata de un requisito previo de la agencia, ya que son límites que, por así decirlo, mantienen a raya al mundo y otorgan al sujeto el distanciamiento que necesita para elegir por sí mismo. Aseguran para el «yo» la capacidad de escoger frente a la vaguedad de las circunstancias que de otra manera lo absorberían. La noción de que los principios de la justicia, como los límites del «yo», aportan una base para la elección que no es en sí elegida, no equivale a una contradicción sino a un presupuesto necesario para un sujeto capaz de elección. A este respecto puede compararse con las nociones paralelas que atraviesan la concepción deontológica, incluyendo al sujeto de la posesión que no es a su vez poseído, una base para el mérito que en sí no es merecida, y un fundamento del contrato que no es contractual.

Hemos visto que la elección de los fines está restringida desde el inicio por los principios de la justicia definidos anteriormente. Pero a fin de describir la explicación de Rawls de la elección y de evaluar el papel de la reflexión, si es que lo tiene, necesitamos saber con mayor detalle la manera en la que se hacen presentes las restricciones de la justicia, la manera exacta en la que se integran a la deliberación del agente. ¿Se incluyen las restricciones de lo justo de alguna manera en la actividad de deliberación, de forma tal que sólo puedan surgir los deseos o concepciones justos del bien, o es el agente el que forma los valores y objetivos basados en ciertos deseos injustos, para suprimirlos en la práctica o apartarlos una vez que resulta claro que violan la justicia?

Por momentos Rawls se expresa como si los principios de la justicia conformaran la concepción del bien de una persona desde el inicio, incluso mientras se formula la concepción. «Al diseñar planes y al decidir sobre sus aspiraciones, los hombres han de tomar en cuenta estas restricciones. [...] [Sus] deseos y aspiraciones se restringen desde el comienzo mediante los principios de la justicia que especifican los límites que los sistemas de fines de los hombres tienen que respetar» (49-50). En otros momentos parece favorecer la segunda explicación,

como por ejemplo cuando afirma que en la justicia como imparcialidad, las personas «convienen [...] implícitamente, en adecuar las concepciones de su bien a lo que requieran los principios de la justicia *o, al menos, en no insistir en pretensiones* que los violen directamente» [Los subrayados son míos.] (49).

Aunque no resulta claro si la justicia interviene en el momento en que elijo mis planes de vida o solamente en un momento posterior, cuando los llevo a cabo, parecería al menos indudable que ninguna de las explicaciones introduce un elemento de autorreflexión del tipo que nos ocupa; ambas conceden que lo justo no determina completamente lo bueno para mí, que mi concepción del bien, aunque restringida dentro de un cierto alcance, aún permanece sujeta a mi elección. En este punto Rawls introduce otra serie de consideraciones para restringir nuestra elección, ciertos «principios contables», como los describe, que en términos generales equivalen a los preceptos básicos de la racionalidad instrumental. Recomienda, por ejemplo, que se elijan los medios más eficaces para lograr fines dados, y no los menos eficaces, un plan más abarcador antes que uno restringido, el plan que ofrezca más posibilidades de éxito y no el que menos, etc. Rawls reconoce, sin embargo, que los principios contables, aun cuando se los complementa con algunos otros principios de elección racional, «no [son] suficientes para asegurar que un proyecto es racional» (466), y que las restricciones de lo justo y de la racionalidad instrumental, tomados en su conjunto, no son suficientes para conducirnos a una única elección determinada. Una vez que se acaban estos principios, simplemente debemos escoger; «podemos reducir el campo de la elección puramente preferente, pero no podemos eliminarlo totalmente. [...] Al fin llegamos a un punto, sin embargo, en el que ya tenemos que decidir qué proyecto preferimos, sin ulterior orientación del principio» (610-609).

Es en este punto, de acuerdo con Rawls, en el que nuestra reflexión debe involucrarse a fin de determinar lo mejor posible qué cosas son las que deseamos y cuánto las deseamos, y determinar a la luz de todos los hechos relevantes el plan que más probablemente realice tales deseos de una forma más completa.

> Voy a suponer que, mientras los principios racionales pueden centrarse en nuestros juicios y establecer líneas orientadoras para la reflexión, al final debemos elegir por nosotros mismos, en

el sentido de que la elección se basa, muchas veces, en nuestro autoconocimiento directo, no sólo de las cosas que deseamos, sino también de la intensidad con que las deseamos. A veces, no es posible evitar la necesidad de valorar la relativa intensidad de nuestros deseos. (460)

Podemos decir que el proyecto racional para una persona es el que elegiría con racionalidad deliberativa (entre los que son compatibles con los principios contables y con otros principios de elección racional, una vez establecidos éstos). Es el proyecto sobre el que recaería la decisión, como resultado de una reflexión cuidadosa, en la que el agente reconsideraría, a la luz de todos los hechos pertinentes, lo que probablemente podría realizar aquellos proyectos, investigando así el modo de acción que mejor cumpliría sus más fundamentales deseos. (461)

La explicación de Rawls de la manera en que efectuamos la elección parecería confirmar el alcance limitado de la reflexión en esta concepción. Mientras se afirma que el plan de vida o concepción del bien más apropiados para una persona en particular son «el resultado de una cuidadosa reflexión», resulta claro que los objetos de tal reflexión están restringidos a: (1) los distintos planes alternativos y sus posibles consecuencias para la realización de los deseos del agente, y (2) los deseos y aspiraciones en sí mismos, y sus intensidades relativas. En ninguno de ambos casos la reflexión toma como su objeto el «yo» en tanto sujeto de deseos. La reflexión que involucra (1), sopesando las alternativas y estimando sus posibles consecuencias, difícilmente puede considerarse como una forma de *auto*rreflexión; se vuelve hacia el exterior y no hacia el interior, y conforma una especie de razonamiento prudencial que en principio podría ser llevado a cabo con igual o mayor éxito por un experto externo que supiera relativamente poco acerca del agente pero que conociera bastante sobre las alternativas en cuestión y los tipos de intereses y deseos que satisfacen típicamente.

La reflexión que implica (2), la evaluación de la intensidad relativa de los deseos, en un sentido vuelca la mirada hacia el interior, pero no *completamente* hacia el interior. Toma como su objeto los deseos, aspiraciones y preferencias contingentes del «yo», pero no al «yo» *en sí mismo*. Dado que para Rawls la facultad de autorreflexión está limitada a la evaluación de la intensidad relativa de los deseos y aspiraciones existentes, la de-

liberación que implica no puede indagar sobre la identidad del agente (¿Quién *soy* yo en realidad?), sino sólo sobre los sentimientos y emociones del agente (¿Qué es lo que realmente *deseo* hacer, o qué *prefiero*?). Debido a que este tipo de deliberación está limitado a la evaluación de los deseos de un sujeto cuya identidad está dada (no reflexivamente) en forma anterior, no puede llevar a la autocomprensión en el sentido fuerte que permite al agente participar en la constitución de su identidad.

Aunque Rawls de hecho se refiere brevemente a «nuestro conocimiento de sí mismo directo» de las cosas que queremos y de cómo las queremos, el «conocimiento de sí mismo» en este sentido no parece mucho más que una conciencia de nuestros deseos y aspiraciones inmediatas. Y en tanto este conocimiento de sí mismo es «directo» en el sentido estricto de que está dado en forma transparente a nuestra comprensión, es difícil imaginar cómo podría tener lugar en cualquier caso algo que se parezca a la reflexión o a la deliberación en el sentido habitual, ya que probablemente sabríamos todo lo que necesitáramos saber «en un instante», antes de que pudiera comenzar ningún proceso reconocible como deliberativo. Incluso si «nuestro conocimiento directo de nosotros mismos» permite una incertidumbre que la reflexión debe resolver, el «yo» que es conocido una vez que la incertidumbre se disipa no es realmente el «yo» en el sentido estricto totalmente distinguido, sino meramente los accidentes y atributos contingentes del «yo».

La diferencia entre el tipo de reflexión que se refiere solamente a los deseos del agente, y el tipo que se extiende al sujeto de los deseos y explora su identidad, corresponde en parte a la distinción realizada por Charles Taylor en su explicación de la facultad de agencia humana entre el «simple evaluador» y el «juez fuerte». Resultan centrales para ambas distinciones las imágenes de superficialidad y profundidad. Para Taylor, el «simple evaluador» es reflexivo en el sentido mínimo de que es capaz de evaluar cursos de acción y de actuar de acuerdo con sus evaluaciones. Pero la reflexión del simple evaluador carece de profundidad en cuanto a que sus evaluaciones están limitadas a un «sentir» inarticulado de las alternativas.

> Mientras que para el simple evaluador lo que está en juego es la conveniencia de las diferentes consumaciones, aquellas definidas por sus deseos *de facto*, para el juez fuerte la reflexión tam-

bién examina los distintos modos de ser del agente posibles. Mientras que una reflexión sobre lo que tenemos más deseos de hacer, que es todo lo que un simple evaluador puede hacer al evaluar motivaciones, nos mantiene por así decirlo en la periferia, una reflexión sobre el tipo de seres que somos nos lleva al centro de nuestra existencia como agentes. [...] En este sentido es más profunda. (Taylor 1977: 114-115)

Para Rawls, la reflexión «acerca del tipo de seres que *somos*», y no sobre el tipo de deseos que *tenemos,* no es una posibilidad, en primer lugar porque el tipo de seres que somos está dado como antecedente y no está sujeto a revisión a la luz de la reflexión o de ninguna otra forma de agencia, y, en segundo lugar, debido a que su «yo» está concebido como desposeído de rasgos constitutivos, investido solamente de atributos contingentes que se mantienen siempre a cierta distancia, y por lo tanto no hay nada *en* el «yo» que la reflexión pueda revisar o aprehender. Para Rawls, la identidad del sujeto nunca puede estar en juego en momentos de elección o de deliberación (a pesar de que sus futuros objetivos y atributos pueden por supuesto verse afectados), ya que los límites que la definen están más allá del alcance de la agencia —tanto voluntarista como cognitiva— que contribuiría a su transformación.

La distinción entre una noción de reflexión como la de Rawls, que está limitada a los objetos de deseo, y una como la de Taylor, que penetra más profundamente para llegar al sujeto de deseo, se corresponde con la distinción entre las concepciones sentimental y constitutiva de comunidad que ya hemos identificado, ya que, en la concepción sentimental, el bien de la comunidad está limitado a los objetivos y sentimientos comunitarios de sujetos previamente individualizados, mientras que, en la concepción constitutiva, se percibe que el bien de la comunidad alcanza a las personas en forma más profunda a fin de describir no sólo sus *sentimientos* sino también un modo de autocomprensión que constituye parcialmente su identidad, que define parcialmente quién es esa persona.

Hemos visto que la teoría de la justicia de Rawls requiere para ser coherente de una concepción de la comunidad en el sentido constitutivo, la cual a su vez exige una noción de la agencia en el sentido cognitivo, y hemos descubierto que su teoría del bien no acepta ninguna de las dos. Esto pone en duda su teoría de la justicia, o su teoría del bien, o ambas.

Pero más allá de las dificultades que surgen para la concepción de Rawls como un todo, sus explicaciones limitadas de la agencia y de la reflexión son poco plausibles en sí mismas, incapaces de comprender en qué pueden consistir la elección y la deliberación. O al menos es esto lo que intentaremos demostrar, considerando la imagen de la elección que nos resta una vez que se ha eliminado la reflexión en el sentido fuerte, general.

La facultad de agencia y el papel de la elección

Como hemos visto, la teoría del bien de Rawls es voluntarista; nuestros objetivos, valores y concepciones del bien fundamentales son materia de nuestra elección, y, al elegirlos, ejercitamos nuestra facultad de agencia. Como lo describe el autor, una vez que se eliminan los principios de la elección racional (es decir, instrumental): «*debemos, al fin, elegir por nosotros mismos*, en el sentido de que la elección se basa, muchas veces, en nuestro autoconocimiento directo, no sólo de las cosas que deseamos, sino también de la intensidad con que las deseamos. [...] *Está claro que se deja al propio agente* la decisión acerca de qué es lo que más desea». [Los subrayados son míos] (460). Dado que los principios de elección racional no especifican un único plan de vida mejor, «muchas cosas quedan todavía por *decidir*. [...] Al fin llegamos a un punto, sin embargo, en el que *ya tenemos que decidir* qué proyecto preferimos, sin ulterior orientación del principio. [...] podemos reducir el campo de *la elección puramente preferente*, pero no podemos eliminarlo totalmente. [...] *Es la propia persona quien debe adoptar esta decisión*, teniendo en cuenta toda la gama de sus inclinaciones y deseos, presentes y futuros» [Los subrayados son míos] (496, 609-610, 616).

Si resulta claro que Rawls describiría mis valores y concepciones del bien como los productos de mi elección o decisión, queda por ver en qué consiste exactamente esta elección y cómo llego yo a realizarla. De acuerdo con su concepción, nosotros elegimos «por nosotros mismos, *en el sentido* de que la elección *se basa, muchas veces*, en nuestro conocimiento directo por nosotros mismos» de lo que deseamos o de cuánto lo de-

seamos. Pero una elección que lo es «en el sentido» de que «se basa, muchas veces en» mis deseos y aspiraciones existentes (¿es determinada por ellos?), resulta una elección sólo en un sentido peculiar del término, ya que suponiendo, como Rawls, que los deseos y aspiraciones en los cuales «se basa» mi elección no son elegidos a su vez sino que son el producto de las circunstancias («no decidimos ahora lo que ahora hemos de desear» 459), una «elección» como ésta implicaría no tanto un acto voluntario como una explicación fáctica de lo que realmente constituyen estos deseos y aspiraciones. Y una vez que yo lograra determinar, por medio del «conocimiento de mí mismo directo», este fragmento de información psicológica, no parecería restar nada para que yo *escoja*. Todavía tendría que coordinar mis deseos y aspiraciones, así determinados, con los mejores medios disponibles de satisfacerlos, pero ésta es una cuestión prudencial que no involucra libre albedrío o ejercicio de voluntad alguno.

Cuando Rawls expresa que «se deja al propio agente la *decisión* acerca de qué es lo que más desea» [Los subrayados son míos] (460), y que «ya tenemos que *decidir* qué proyecto preferimos» (609), la «decisión» que debe realizar el agente equivale nada más que a una estimación o inventario psicológico de los deseos y preferencias que ya tiene, no a una elección de los valores que profesaría o de los objetivos que procuraría. Como en el caso de la «elección» o «acuerdo» colectivos de la posición original, tal «decisión» no decide nada más que cuán correctamente ha percibido el agente algo que ya existía, en este caso la forma y la intensidad de estos deseos preexistentes. Pero, en este caso, el aspecto voluntarista de la agencia parecería desvanecerse por completo.

Llegar a un plan de vida o una concepción del bien simplemente obedeciendo a mis deseos y aspiraciones existentes equivale a no elegir ni el plan ni los deseos; simplemente equivale a coordinar los objetivos que ya tengo con los mejores medios disponibles para satisfacerlos. De acuerdo con esta descripción, mis objetivos, valores y concepciones del bien no son producto de la elección sino los objetos de una determinada introspección superficial, lo suficientemente «interior» como para estudiar de manera no crítica los motivos y deseos que los accidentes de mis circunstancias me han dejado; yo simplemente los conozco de la misma manera que los siento y busco mi camino lo mejor posible hacia su consumación.

Puede sugerirse que Rawls podría evitar el colapso aparente de esta explicación de la agencia y de la elección de alguna de las siguientes maneras. La primera consistiría en introducir la idea de que las personas son capaces de reflexionar sobre sus deseos no sólo en el sentido de evaluar cuán intensos son sino también en el sentido de evaluar qué deseables le resultan; capaces, en suma, de formar deseos de segundo orden, cuyos objetos sean ciertos deseos de primer orden (Frankfurt 1971). Por lo tanto, puedo querer tener ciertos deseos y no otros, o considerar ciertos tipos de deseos como convenientes y otros no tanto. El hecho de que algo sea deseado (y no sea injusto) no sería ya suficiente para volverlo bueno, ya que esto dependería de la cuestión ulterior de si se trata de un deseo conveniente o no. Una vez que yo determinara qué es lo que deseo (realmente) en términos de deseo de primer orden, me restaría evaluar cuán conveniente es mi deseo y en este sentido afirmarlo o rechazarlo.

De hecho, Rawls parece admitir vagamente tal posibilidad cuando afirma que, aunque «no decidimos ahora lo que ahora hemos de desear», podemos al menos «decidir ahora hacer algo que sabemos que influirá en los deseos que tendremos en el futuro. [...] Así, elegimos entre deseos futuros, a la luz de nuestros deseos presentes» (459).

Pero incluso si un agente en el sentido de Rawls fuera capaz de formar deseos con respecto a ciertos deseos distintos, su agencia no sería restaurada en ningún sentido significativo, ya que no tendría fundamentos con los cuales justificar o defender que un tipo de deseo sea más conveniente o deseable que otro, más allá del mero hecho de su deseo de segundo orden. Todavía podría apelar solamente al hecho psicológico de su preferencia (ahora de segundo orden), y sólo tendría su intensidad relativa para evaluar. Ni el valor intrínseco de un deseo ni su conexión esencial con la identidad del agente pueden aportar una base para afirmarlo, ya que en la concepción de Rawls, el valor de un deseo aparece sólo a la luz del bien de una persona, y la identidad del agente está despojada de rasgos constitutivos de manera tal que ningún objetivo ni deseo puede resultar esencial para ella. La afirmación o rechazo de los deseos sugeridos por la formación de los deseos de segundo orden no introduciría en los supuestos de Rawls un ulterior elemento de reflexión o voluntad, ya que tal evaluación sólo reflejaría una estimación ligeramente más compleja de la in-

tensidad relativa de los deseos preexistentes, incluyendo los de primer y segundo orden. No podría decirse que la concepción del bien resultante fue elegida, en el mismo sentido que no lo sería la que surgiera sólo de deseos de primer orden.

Un segundo intento para restablecer la coherencia de la elección en la concepción de Rawls podría consistir en imaginar un caso en el cual los diversos deseos del agente, evaluados adecuadamente en cuanto a sus respectivas intensidades, llevarían a un empate, y donde la deliberación ya hubiera tomado en cuenta todas las preferencias relevantes de forma tal que no pudiera introducirse nueva preferencia alguna para romper dicho empate. En este caso, podría continuar esta explicación, el agente no tendría más alternativa que inclinar la balanza, en forma arbitraria, hacia uno u otro lado, sin basarse en ninguna preferencia ni deseo en absoluto. Podría sugerirse que una «elección» independiente de la influencia de deseos y aspiraciones preexistentes —una «elección radicalmente libre», como se la describe en ocasiones— admitiría el aspecto voluntarista de la agencia que parece no estar presente cuando el agente está obligado a «elegir» en conformidad con deseos y aspiraciones preexistentes.

Pero Rawls rechaza la idea de una forma de agencia totalmente arbitraria que podría eludir la influencia de deseos y aspiraciones preexistentes. «La noción de la elección radical [...] no tiene lugar alguno en la justicia como imparcialidad» (Rawls 1980: 568). A diferencia de los principios de lo justo, que expresan la autonomía del agente y deben encontrarse libres de las contingencias, las concepciones de lo bueno se entienden como completamente heterónomas. Cuando surgen objetivos incompatibles, Rawls no habla de una elección arbitraria o radicalmente libre, sino de una «elección puramente preferente», sugiriendo la forma de (no) agencia que consideramos en primer lugar. De todos modos, la noción de una «elección» puramente arbitraria no regida por consideración alguna es difícilmente una explicación de la agencia voluntarista más plausible que la de la «elección» guiada totalmente por las preferencias y deseos predeterminados. Ni la «elección puramente preferente» ni la «elección puramente arbitraria» pueden redimir la noción de agencia de Rawls en el sentido voluntarista; la primera confunde elección con necesidad, la segunda con capricho. Juntas reflejan el alcance limitado de la reflexión dentro de la concepción de Rawls, y la poco plausible explicación de la agencia humana que se deriva de ella.

La categoría del bien

La dificultad que plantea la teoría del bien de Rawls es epistemológica tanto como moral, y en esto nos recuerda un problema que ha surgido en relación con el concepto de lo justo; el de distinguir categorías de evaluación a partir de lo que se está evaluando. Si mis valores fundamentales y mis fines últimos me permiten, como deben hacerlo, evaluar y regular mis deseos y aspiraciones inmediatas, tales valores y fines deben tener una justificación independiente del mero hecho de que casualmente los tenga con cierta intensidad. Pero si mi concepción del bien es simplemente el producto de mis deseos y aspiraciones inmediatos, no existe razón para suponer que el punto de partida crítico que aporta tenga más valor que los deseos que trata de evaluar; en tanto producto de esos deseos, sólo se regiría por las mismas contingencias.

Rawls responde a esta dificultad en el caso de lo justo buscando un punto de Arquímedes en la justicia como imparcialidad que «no está a merced de los deseos e intereses presentes» (298). Pero tal como hemos visto, el concepto de lo justo de Rawls no se extiende a la moralidad privada, ni ningún otro instrumento de distanciamiento protege al bien de estar cabalmente implicado en los deseos e intereses presentes del agente. La «elección puramente preferente» es una elección completamente heterónoma, y los valores o concepciones del bien de una persona pueden llegar más allá de esto. Como acepta Rawls, para nuestra sorpresa: «El que tengamos una concepción del bien en lugar de otra no es relevante desde el punto de vista moral. Al adquirirla hemos estado influenciados por el mismo tipo de contingencias que nos llevaron a eliminar el conocimiento de nuestro sexo y clase social» (1975: 537).

El limitado alcance dado a la reflexión dentro de la concepción de Rawls, y la problemática, e incluso debilitada, teoría del bien que surge como resultado, revelan la dimensión en la que el liberalismo deontológico acepta una explicación esencialmente utilitarista del bien, a pesar de que pueda diferir su teoría del bien. Este entorno utilitarista apareció por primera vez en nuestro debate sobre la defensa de Dworkin de la acción afirmativa; una vez que se considera que no hay derechos individuales en juego, prevalecen automáticamente las consi-

deraciones utilitaristas. Aunque Dworkin defiende lo que él llama un «concepto de lo justo contrario al utilitarismo», el alcance de lo justo está estrictamente circunscrito (aunque en forma elusiva), de forma tal que «el grueso de las leyes que disminuyen mi libertad están justificadas sobre bases utilitaristas, en el sentido de que sirven al interés general o al bienestar común» (Dworkin 1977a: 269).[17]

Los antecedentes utilitaristas de la concepción de Rawls se revelan más claramente en las referencias a la vida moral individual. Mientras que la justicia como imparcialidad rechaza al utilitarismo como base de la moralidad social o pública, no tiene aparentemente ningún argumento en contra del utilitarismo como base de la moral individual o privada, al contrario de la noción kantiana del «deber hacia sí mismo». Rawls describe la explicación utilitarista de la moral privada, sin objeción discernible, de la siguiente manera:

> Una persona actúa de manera correcta, al menos cuando otros no resultan afectados, cuando trata de obtener el mayor beneficio posible y de promover sus fines racionales. [...] [E]l principio para un individuo es promover tanto como sea posible su propio bienestar, esto es, su propio sistema de deseos [...] (41/2).

> Desde luego, hay un principio formal que parece facilitar una respuesta general [a la elección del plan de vida de un individuo]. Éste es el principio de adoptar el proyecto que arroje el más elevado saldo de satisfacciones. (460)

Para Rawls, el error del utilitarismo no radica en su concepción del bien como la satisfacción de deseos dados arbitrariamente y no diferenciados en cuanto a su valor —ya que la justicia como imparcialidad comparte esto—, sino sólo en su indiferencia hacia la manera en que estas realizaciones se reparten entre los individuos. Su error, para él, yace en adoptar «para la sociedad, como un todo, el principio de elección racional del individuo», el combinar «los deseos de todas las personas en un sistema coherente de deseos», y buscar su satisfacción general (45). Al hacerlo, «fusiona» o «funde» a todas las personas en una sola, reduce la elección social a «esencialmente una cuestión de administración eficiente» (como, pode-

17. Véase la obra de H.L.A. Hart (1979: 86-89), donde se encontrará una crítica muy interesante de la perspectiva de Dworkin sobre este tema.

mos presumir, puede reducirse adecuadamente la elección individual) y por lo tanto no logra tomar seriamente en cuenta la distinción entre las personas (47, 53).

La justicia como imparcialidad trata de solucionar estos fallos al enfatizar la distinción entre las personas y al insistir en la separación entre los diversos «sistemas de deseos» que el utilitarismo confunde. Pero no resulta inmediatamente evidente el fundamento de la divergencia entre Rawls y el utilitarismo a este respecto. Aunque el primero parece creer firmemente en su visión de que a cada ser humano individual le corresponde exactamente un «sistema de deseos», nunca establece por qué esto debe ser así, o qué es exactamente un «sistema de deseos», o por qué es incorrecto confundirlos en uno solo. ¿Un «sistema de deseos» es un grupo de deseos *ordenado* de una cierta manera, agrupados en una jerarquía de valor relativo o conexión esencial con la identidad del agente, o es simplemente una cadena de deseos arbitrariamente reunidos, que sólo se distinguen por su intensidad relativa y su ubicación accidental? En el segundo caso, si un *sistema* de deseos significa nada más que una colección arbitraria de deseos accidentalmente incorporados a un ser humano en particular, no resulta muy claro por qué la integridad de tal «sistema» debería tomarse tan seriamente desde lo metafísico y lo moral. Si los deseos pueden confundirse adecuadamente dentro de las personas, ¿por qué no también entre las personas?

Si, por otra parte, lo que lo vuelve un *sistema* de deseos es un orden jerárquico de deseos distinguibles cualitativamente, no parecería más justificable el «confundir» los deseos dentro de una persona que entre personas distintas, y el error del utilitarismo también sería un error, al menos a este respecto, de la justicia como imparcialidad. La tendencia a confundir los deseos, ya sea dentro de una persona o entre personas distintas, reflejaría un fallo en su *ordenación*, o en reconocer las distinciones cualitativas entre ellos. Pero este fallo atraviesa la distinción entre la elección individual y la social, ya que no hay razón para suponer que un «sistema de deseos» en *este* sentido corresponda en todos los casos a una persona empíricamente individualizada. Las comunidades de tipos distintos contarían como «sistemas de deseos» distintos en este sentido, en tanto fueran identificables en parte por un orden o estructura de valores compartidos parcialmente constitutivo de una identidad o forma de vida común. Desde este punto de vista, el

fracaso utilitarista al no lograr tomar en serio la distinción entre las personas aparecería como un mero síntoma de este fracaso más significativo en tomar seriamente en cuenta las distinciones cualitativas de valor entre los diferentes órdenes de deseos, un error que tiene sus raíces en una concepción empobrecida del bien que, como se ha visto, comparte la justicia como imparcialidad.

Podría pensarse que para una doctrina deontológica como la de Rawls la percepción del bien como completamente inmerso en la contingencia, a pesar de su poca plausibilidad general, tendría al menos la ventaja redentora de hacer más apremiante la primacía de lo justo. Si el bien no resulta otra cosa que la satisfacción indiscriminada de preferencias dadas arbitrariamente, sin relación con su valor, no es difícil imaginar que lo justo (y en todo caso una cierta cantidad de otros tipos de demandas) deba tener más peso. Pero de hecho la categoría moralmente reducida del bien debe inevitablemente cuestionar también la categoría de la justicia, ya que una vez que se admite que nuestras concepciones de lo bueno son moralmente arbitrarias, resulta difícil entender por qué la más alta de todas las virtudes (sociales) debe ser la que nos permita perseguir estas concepciones arbitrarias tan plenamente como lo permitan las circunstancias.

La epistemología moral de la justicia

Nuestro estudio del bien nos lleva así nuevamente a la cuestión de la justicia y al postulado de su prioridad, y con esto regresamos a las circunstancias de la justicia en la posición original. Aquí, la distinción o separación entre las personas, sobre la cual Rawls insiste como una corrección al utilitarismo, se instala como el supuesto clave del desinterés mutuo, la noción de que los individuos no se preocupan por los intereses de los demás (153-154). Cuando consideramos por primera vez las condiciones en la posición original, este supuesto en particular y la justificación empirista de las circunstancias de la justicia en general parecían socavar la primacía de la justicia por diversas razones. En el caso de que la justicia dependiera para su virtud de la existencia de ciertas

condiciones empíricas previas, la virtud de la justicia no sería ya absoluta, como la verdad lo es para las teorías, sino solamente condicional, como el coraje físico lo es en una zona de guerra; presupondría una virtud o grupo de virtudes rivales con al menos una categoría correlativa, asumiría en ciertas circunstancias una dimensión compensatoria, y finalmente, cuando se utilizara en forma inadecuada, la justicia aparecería como un vicio en lugar de una virtud. En suma, una concepción similar a la de Hume de las circunstancias de la justicia —como la que adopta explícitamente Rawls— parecería incompatible con la categoría privilegiada que Rawls requiere para la justicia y que Kant defiende sólo como recurso a una metafísica moral que el primero consideraba inaceptable.

La perspectiva propia de Hume sobre la justicia confirma su parcialidad, al menos en cuanto a que se deriva de premisas que Hume y Rawls parecen compartir. Para Hume, las circunstancias de la justicia describen ciertas condiciones materiales y de motivación desafortunadas, aunque inevitables, de las sociedades humanas reales, especialmente la escasez moderada y una «generosidad limitada». Reunidas, estas circunstancias demuestran en qué sentido la llegada de la justicia significa la ausencia de ciertas virtudes más nobles, aunque más raras.

«Si cada hombre tuviera un tierno aprecio por el prójimo, o si la naturaleza proveyera abundantemente de todas nuestras necesidades y deseos, [...] los celos del interés, que la justicia supone, ya no podrían tener lugar», ni, afirma Hume, habría ocasión para la distinción entre propiedad y posesión. «Elévense en grado suficiente la benevolencia de los hombres o el botín de la naturaleza y haréis que la justicia se convierta en algo inútil, supliendo su lugar mediante virtudes mucho más nobles y bienes más valiosos». Si la escasez natural fuera reemplazada por la abundancia, «o si todos tuvieran el mismo afecto y tierno aprecio por los demás como por sí mismos, la justicia y la injusticia serían igualmente desconocidas entre los seres humanos». Y así, concluye Hume, «es sólo debido a las escasas provisiones que la naturaleza ha hecho para nuestras necesidades, que la justicia deriva su origen» (Hume 1739: 494-5).

Para Hume, la justicia no puede ser la primera virtud de las instituciones sociales (al menos no en ningún sentido categórico), y en algunos casos resulta dudoso si se trata de una virtud. En la institución de la familia, por ejemplo, los afectos pueden

estar ampliados en tal medida que la justicia apenas se vea implicada, mucho menos como «la primera virtud». E incluso en la sociedad más amplia, donde la generosidad está más limitada y la justicia se ve más frecuentemente implicada, su virtud sólo puede explicarse contra un marco de virtudes más altas o nobles cuya ausencia crea la necesidad de justicia. En tanto pudieran cultivarse en mayor medida la benevolencia mutua y los afectos ampliados, la necesidad de «la cautelosa, celosa virtud de la justicia» disminuiría en proporción, y la humanidad se encontraría mejor por ello. En el caso de que la escasez o el egoísmo estuvieran totalmente superados, «la justicia, al ser completamente inútil [...] nunca podría tener un lugar en el catálogo de las virtudes» (Hume 1777: 16), y mucho menos el primer lugar que Rawls le asignaría.

Pero a pesar del paralelismo entre la concepción de Hume y la suya, al que invita el mismo Rawls, el supuesto del desinterés mutuo tiene para este último un significado diferente. No implica que los seres humanos están gobernados típicamente por «el egoísmo y la generosidad restringida»; de hecho no se lo propone como un postulado sobre las motivaciones humanas en absoluto, sino como un postulado sobre el sujeto de las motivaciones. Supone intereses *de* un «yo», no necesariamente *en* un «yo»: un sujeto de la posesión individualizado por adelantado y dado con anterioridad a sus fines.

De esto se desprenden importantes consecuencias para la categoría de la justicia. La benevolencia ya no es anterior a la justicia y en algunos casos capaz de reemplazarla. Dado que para Rawls la virtud de la justicia no presupone motivaciones egoístas desde el principio, no necesita esperar el derrumbe de la benevolencia para encontrar su ocasión de aparecer, e incluso el completo florecimiento de los «afectos ampliados» no puede desplazarla. La justicia deja de ser meramente compensatoria con respecto a las «virtudes más nobles», ya que su virtud ya no depende de su ausencia. Por el contrario, cuando las personas están individualizadas en el sentido de Rawls, la justicia no sólo gana su independencia de los sentimientos y motivaciones que prevalecen, sino que se ubica por encima de ellos como primaria. Dada la naturaleza del sujeto tal como lo concibe Rawls, la justicia no es meramente un sentimiento o una emoción como otras virtudes, menores, sino por encima de todo un marco que restringe tales virtudes y es «regulativo» con respecto a ellas.

> Por tanto, para realizar nuestra naturaleza no tenemos más alternativa que la de proteger nuestro sentido de la justicia para que rija nuestros restantes objetivos. Este sentimiento no puede realizarse, si está comprometido y equilibrado frente a otros fines, como un solo deseo entre los demás. [...] Por el contrario, la medida en que logramos expresar nuestra naturaleza depende del grado de coherencia con que actuamos, a partir de nuestro sentido de la justicia como regulador último. Lo que no podemos hacer es expresar nuestra naturaleza siguiendo un proyecto que considera el sentido de la justicia sólo como un deseo que ha de valorarse frente a otros. Porque este sentimiento revela lo que la persona es, y transigir en él equivale a no alcanzar para el «yo» el reino de la libertad, sino dar paso a las contingencias y a los accidentes del mundo. (634/35)

Hemos visto de qué manera la prioridad de la justicia, como la prioridad del «yo», se deriva en gran parte de su libertad respecto de las contingencias y accidentes del mundo. Esto quedó establecido en nuestra argumentación sobre lo justo y los límites del «yo». A la luz de nuestro debate acerca del bien, podemos ahora ver por qué, en la teoría del sujeto de Rawls, virtudes tales como la benevolencia e incluso el amor no son ideales morales autosuficientes sino que deben aguardar a la justicia para estar completos.

Dado el limitado papel que juega la reflexión en la explicación de Rawls, las virtudes de la benevolencia y del amor, como rasgos del bien, son formas de sentimiento más que de conocimiento, maneras de sentir y no de saber. A diferencia de las emociones o sentimientos personales o de primer orden, cuyos objetos están dados más o menos directamente por mi conciencia, la benevolencia y el amor son deseos cuyo objeto es el bien de otro. Pero dada la separación de las personas y la inflexibilidad de los límites entre ellas, el contenido de *este* bien (es decir, el bien que yo le deseo a otro) debe resultarme en gran medida opaco. En la perspectiva de Rawls, el amor es ciego, no por su intensidad sino por la opacidad del bien que es el objeto de esta preocupación. «La razón por la cual la situación permanece oscura es que el amor y la benevolencia son nociones de segundo orden: intentan promover el bien previamente existente de los individuos amados» (222).

Si el llegar a nuestro propio bien es principalmente una cuestión de investigar las preferencias presentes y evaluar sus intensidades relativas, no resulta el tipo de investigación

en la cual otro, incluso otro con quien se tenga una íntima relación, pueda participar fácilmente. Sólo la persona misma puede «saber» lo que realmente desea o «decidir» lo que prefiere. «Aún cuando adoptemos el punto de vista de otro e intentemos apreciar lo que sería conveniente para él, lo hacemos como un consejero, por así decirlo» (495), y dado el limitado acceso cognitivo que admite la concepción de Rawls, en todo caso un consejero bastante poco dotado.

Aunque algunas veces podemos superar la dificultad de conocer el bien de un individuo amado cuyos intereses procuraríamos, el problema se vuelve irremisiblemente complejo si extendiéramos nuestro amor o benevolencia a una pluralidad de personas cuyos intereses pueden estar en conflicto. Esto se debe a que no podemos esperar conocer sus respectivos bienes lo suficiente como para diferenciarlos y evaluar sus demandas relativas. Incluso aunque la benevolencia pudiera cultivarse tan ampliamente como sugiere Hume en su visión hipotética, su virtud no sería aún autosuficiente, ya que seguiría resultando poco claro, sin más, qué es lo que nos dictaría el amor a la humanidad. «De nada sirve decir que hemos de juzgar la situación tal y como lo dicta la benevolencia. Eso presupone que estaríamos orientados equivocadamente por el interés propio. El problema radica en otra parte. La benevolencia fracasa en la medida en que sus muchos amores entran en oposición en las personas que constituyen sus múltiples objetos» (221). No resulta sorprendente que el ancla que requiere esta benevolencia esté dada por la virtud de la justicia; la benevolencia, incluso en la interpretación más expansiva, depende de la justicia para estar completa. «[U]n amor a la humanidad que desee preservar la distinción entre las personas, para reconocer como separadas sus vidas y sus experiencias, usará los dos principios de la justicia para determinar sus objetivos, cuando los muchos bienes que desea servir se encuentren en oposición» (222). Incluso frente a una virtud tan noble como el amor por la humanidad, prevalece la primacía de la justicia, aunque el amor que resulta tenga un espíritu extrañamente judicial.

> Este amor está guiado por lo que los individuos mismos acordarían en una situación inicial equitativa que les proporcionara una representación igual en tanto que personas morales. (222)

Así, vemos que el supuesto del desinterés mutuo de las partes no impide una interpretación razonable de la benevolencia y del amor a la humanidad *dentro del marco de la justicia como imparcialidad*. [Los subrayados son míos.] (222)

Para Rawls, las consecuencias de tener seriamente en cuenta la distinción entre las personas no son directamente morales sino más decisivamente epistemológicas. Lo que los límites entre las personas confinan no es tanto el alcance de nuestros sentimientos —ya que no los prejuzgan— sino el alcance de nuestra comprensión, nuestro acceso cognitivo a los demás. Y es este déficit *epistémico* (que se deriva de la naturaleza del sujeto) más que cualquier carencia de benevolencia (que en todo caso es variable y contingente) lo que exige de la justicia para ser compensado, y lo que justifica su preeminencia. Aunque para Hume necesitamos la justicia debido a que no nos *amamos* lo suficiente, para Rawls necesitamos la justicia porque no podemos *conocernos* lo suficiente como para que incluso el amor sirva por sí mismo.

No obstante, como sugiere nuestra argumentación sobre la agencia y la reflexión, no somos ni tan transparentes para nosotros mismos ni tan opacos para los demás como afirma la epistemología moral de Rawls. Si nuestra agencia debe consistir en algo más que el ejercicio de la «administración eficiente» que implica la concepción de Rawls, debemos ser capaces de una introspección más profunda que la que admite un «conocimiento propio directo» de nuestros deseos y aspiraciones inmediatas. Pero para ser capaces de una reflexión más profunda, no podemos ser sujetos de la posesión completamente descarnados, individualizados anticipadamente y dados con anterioridad a nuestros fines, sino que debemos ser sujetos constituidos en parte por nuestras aspiraciones y vinculaciones centrales, siempre abiertos, y de hecho vulnerables al crecimiento y la transformación a la luz de la revisión de nuestro conocimiento propio. Y en tanto nuestra autocomprensión constitutiva comprenda un sujeto más amplio que el mero individuo, ya sea una familia, una tribu, una ciudad, una clase social, una nación o un pueblo, en esta medida definirá una comunidad en el sentido constitutivo. Y lo que distingue a dicha comunidad no es meramente un espíritu de benevolencia, o el que prevalezcan valores comunitarios, o incluso ciertos «fines últimos compartidos» por sí, sino un voca-

bulario de discurso común y antecedentes de prácticas y entendimientos implícitos dentro de los cuales la opacidad de los participantes se atenúa, aunque nunca se llegue a disolver. En tanto la justicia dependa para su preeminencia de la separación o limitación de las personas en el sentido cognitivo, su prioridad disminuye a medida que se disuelve la opacidad y se profundiza en la comunidad.

La justicia y la comunidad

Siempre podemos preguntarnos, sobre cualquier sociedad, hasta qué punto es justa, o «bien ordenada» en el sentido de Rawls, y hasta qué punto es una comunidad. La respuesta no puede en ninguno de los dos casos estar dada completamente por referencia a los sentimientos y deseos de los participantes por sí solos. Como observa el autor, preguntarnos si una sociedad en particular es justa no equivale simplemente a preguntarnos si un gran número de sus miembros tienen por casualidad entre sus diversos deseos el de actuar en forma justa —aunque éste puede ser un rasgo de una sociedad justa—, sino preguntarnos si la sociedad es en sí una sociedad de un cierto tipo, ordenada de cierta manera tal que la justicia describa su «estructura básica» y no meramente las disposiciones de las personas dentro de tal estructura. Así, Rawls afirma que aunque denominemos a las actitudes y disposiciones de las personas como justas o injustas, para la justicia como imparcialidad «el objeto primario de la justicia es la estructura básica de la sociedad» (23). Para que una sociedad resulte justa en este sentido fuerte, la justicia debe ser constitutiva de su marco estructural y no simplemente un atributo de los planes de vida de ciertos participantes.

De manera similar, preguntarnos si una sociedad en particular constituye una comunidad no implica simplemente preguntarnos si una gran cantidad de sus miembros tienen casualmente entre sus diversos deseos el de asociarse con otros o promover objetivos comunitarios —aunque éste puede ser uno de los rasgos de una comunidad—, sino si la sociedad es en sí una sociedad de cierto tipo, ordenada de una cierta manera, de forma tal que la comunidad describa su estructura

básica y no meramente las disposiciones de las personas dentro de la estructura. Para que una sociedad constituya una comunidad en este sentido fuerte, la comunidad debe ser constitutiva de la autocomprensión compartida de los participantes, y estar incorporada en sus acuerdos institucionales, y no ser simplemente un atributo de los planes de vida de ciertos participantes.

Rawls podría alegar que debería rechazarse una concepción constitutiva de la comunidad como ésta «por razones de claridad entre otras», o sobre la base de que supone que la sociedad es «un todo orgánico con una vida propia, distinta y superior a la de sus miembros en sus relaciones con los demás» (301). Pero, desde lo metafísico, una concepción constitutiva de la comunidad no es más problemática que una concepción constitutiva de la justicia como la que defiende Rawls, ya que si esta noción de comunidad describe el marco de autocomprensión que es distinguible de los sentimientos y disposiciones de los individuos dentro del referido marco, y es en cierto sentido anterior a ellos, lo es sólo en el mismo sentido en que la justicia como imparcialidad describe una «estructura básica» o marco que es así distinguible de los sentimientos y disposiciones de los individuos dentro de tal estructura, y anterior a ellos.

Si el utilitarismo no logra tomar seriamente en cuenta la distinción entre las personas, la justicia como imparcialidad no logra hacer lo propio con respecto a nuestros rasgos comunes. Al considerar los límites del «yo» como anteriores, fijos de una vez y para siempre, confina nuestros aspectos comunes a un aspecto del bien, y relega el bien al nivel de una mera contingencia, un producto de deseos y aspiraciones indiscriminados «no [...] relevantes desde un punto de vista moral». Dada una concepción del bien así reducida, la prioridad de lo justo parecería un postulado ciertamente ineludible. Pero el utilitarismo ha dado al bien un cierto mal nombre, y al adoptarlo en forma poco crítica, la justicia como imparcialidad gana una falsa victoria para la deontología.

Conclusión

El liberalismo y los límites de la justicia

Para que la justicia resulte la primera virtud, deben cumplirse ciertos supuestos acerca de nosotros mismos. Debemos ser criaturas de cierto tipo, relacionarnos con las circunstancias humanas de cierta manera. Debemos mantener una cierta distancia con respecto a nuestras circunstancias, ya sea como sujetos trascendentes, en el caso de Kant, o como sujetos de la posesión esencialmente no restringidos, en el caso de Rawls. De cualquier manera, debemos considerarnos independientes; independientes de los intereses y los lazos que podamos tener en cualquier momento, nunca identificados por nuestros objetivos sino siempre capaces de distanciarnos para investigar y considerar, y posiblemente también evaluar, tales objetivos (Rawls 1979: 7, 1980: 544-545).

El proyecto liberador de la deontología

La visión del universo moral que el «yo» debe habitar está vinculada a la noción de un «yo» independiente. A diferencia de la concepción griega clásica y de la cristiana medieval, el universo de la ética deontológica es un lugar despojado de significado inherente, un mundo «desencantado» según la frase de Max Weber, un mundo sin un orden moral objetivo. Sólo en

un universo desprovisto de *telos*, como el que afirmaban la ciencia y la filosofía del siglo XVII,[18] es posible concebir un sujeto apartado de sus propósitos y fines, y anterior a éstos. Sólo un mundo no gobernado por un orden deliberado deja los principios de la justicia abiertos a la interpretación humana, y somete las concepciones del bien a la elección individual. Es aquí donde aparece más completamente la profundidad de la oposición entre el liberalismo deontológico y las concepciones teleológicas del mundo.

En los casos en que ni la naturaleza ni el cosmos proporcionan un orden significativo que deba ser aprehendido o comprendido, queda a los sujetos humanos el constituir un significado propio. Esto explicaría la preeminencia de la teoría del contrato desde Hobbes en adelante, y el correspondiente énfasis sobre la ética voluntarista, opuesta a la cognitiva, que culmina con Kant. Lo que ya no puede encontrarse queda entonces por crearse.[19] Rawls describe su propia perspectiva al respecto como una versión del «constructivismo» kantiano.

> Las partes en la posición original no están de acuerdo en cuáles son los hechos morales, como si ya existiesen tales hechos. No es que, al estar situadas imparcialmente, tengan una visión clara y no distorsionada de un orden moral anterior e independiente. En lugar de esto (para el constructivismo) *no existe un orden tal*, y por lo tanto tampoco tales hechos aparte del procedimiento como un todo. [Los subrayados son míos.] (1980: 568)

En forma similar, para Kant, la ley moral no es un descubrimiento de la razón teórica sino un fruto de la razón práctica, el producto de la voluntad pura. «Los conceptos prácticos elementales tienen como su fundamento la forma de una voluntad pura dada en la razón», y lo que presta autoridad a

18. Acerca de las consecuencias morales, políticas y epistemológicas de la revolución científica y de las perspectivas del mundo del siglo XVII, véase Strauss 1953; Arendt 1958: 248-325; Wolin 1960: 239-285; y Taylor 1975: 3-50.

19. Como afirma audazmente un liberal, «La dura verdad es ésta: no existe un significado moral oculto en las entrañas del universo [...] No hay necesidad de sentirse agobiados por el vacío. Podemos crear nuestros propios significados, usted y yo». (Ackerman, 1980: 368). Aunque parezca extraño, insiste sin embargo en que el liberalismo no está comprometido con ninguna metafísica o epistemología en particular ni con ninguna «gran pregunta de un carácter altamente comprometido» (356-357, 361).

esta voluntad es el hecho de que legisla en un mundo en donde aún no ha llegado el significado. La razón práctica tiene una ventaja sobre la razón teórica precisamente en esta facultad voluntarista, en su capacidad de generar preceptos prácticos directamente, sin recurso a la cognición. «Dado que en todos los preceptos de la voluntad pura es sólo una cuestión de la determinación de la voluntad», no hay necesidad de que esos preceptos «deban esperar a las intuiciones para adquirir un significado. Esto se da por la notable razón de que *en sí mismos producen la realidad de aquello a lo que se refieren*» [Los subrayados son míos] (Kant 1788: 176).

Es importante tener en cuenta que en la perspectiva deontológica la noción de un «yo» despojado de objetivos y lazos esenciales no implica que seamos seres totalmente carentes de propósitos o incapaces de vínculos morales, sino que los valores y relaciones que tenemos son el producto de la elección, ya que la posesión de un «yo» es dada anteriormente a los fines. En el universo deontológico se da de forma similar. A pesar de que rechaza la posibilidad de un orden moral objetivo, este liberalismo no sostiene un «vale todo». Afirma la justicia, no el nihilismo. La noción de un universo vacío de significado intrínseco no implica, en la perspectiva deontológica, un mundo donde los principios regulativos no rijan en absoluto, sino un universo moral habitado por sujetos capaces de constituir un significado propio, como agentes de *construcción* en el caso de lo justo, y como agentes de *elección* en el caso del bien. En tanto «yo» noumenal, o en tanto parte en la posición original, puedo llegar a los principios de la justicia; en tanto «yo» individual y real, puedo llegar a las concepciones de lo bueno. Y los principios que construya como un «yo» noumenal restringen (pero no determinan) los propósitos que elijo como «yo» individual. Esto refleja la prioridad de lo justo sobre lo bueno.

El universo deontológico y el «yo» independiente que se mueve dentro de él, tomados en su conjunto, sostienen una visión liberadora. Liberado de los dictados de la naturaleza y de la sanción de los roles sociales, el sujeto deontológico está instalado como soberano, ubicado como el autor de los únicos significados morales que existen. Como habitantes de un mundo sin *telos*, somos libres de construir principios de la justicia no restringidos por un orden de valores que esté dado anteriormente. Aunque los principios de la justicia no son en sentido

estricto una cuestión de elección, la sociedad que definen «se acerca, tanto como puede la sociedad, a ser un esquema voluntario» (13), ya que surgen de la voluntad pura o de un acto de interpretación que no responde a un orden moral anterior. En tanto un «yo» independiente, soy libre para elegir mis propósitos y fines sin verme restringido por este orden, o por la costumbre, la tradición o una condición hereditaria. En tanto no sean injustas, nuestras concepciones del bien tienen un valor, sean cuales fueran, simplemente en virtud de que las hayamos escogido. Estamos «originando por nosotros mismos fuentes de demandas válidas» (Rawls 1980: 543).

La justicia es la virtud que personifica la visión liberadora de la deontología y le permite desenvolverse. Incorpora esta visión al describir aquellos principios que el sujeto soberano construye mientras está situado anteriormente a la constitución de todo valor. Permite que la visión se desenvuelva debido a que, equipada con estos principios, la sociedad justa regula la elección de fines de cada persona de una manera compatible con una libertad similar para todos. Los ciudadanos gobernados por la justicia pueden entonces realizar el proyecto liberador de la deontología —el ejercicio de su capacidad como «fuentes [autónomas] de demandas válidas»— en forma tan completa como las circunstancias lo permitan. De esta forma la primacía de la justicia expresa, a un tiempo que anima a las aspiraciones liberadoras de una perspectiva deontológica del mundo y de una concepción del «yo».

Pero la visión deontológica tiene una falla, tanto en sus propios términos como en términos generales en tanto explicación de nuestra experiencia moral. En sus propios términos, el «yo» deontológico, despojado de todos los lazos constitutivos posibles, no es un «yo» liberado sino un «yo» privado de poder. Como hemos visto, ni lo justo ni lo bueno admiten la derivación voluntarista que requiere la deontología. Como agentes de construcción, no construimos realmente (capítulo 3), y, como agentes de elección, no elegimos realmente (capítulo 4). Lo que sucede detrás del velo de la ignorancia no es un contrato o un acuerdo sino un tipo de descubrimiento; y lo que sucede en la «elección puramente preferente» no es tanto una elección de fines como una correlación entre deseos preexistentes, no diferenciados en cuanto a su valor, y los mejores medios disponibles para satisfacerlos. Para las partes en la posición original, al igual que para las partes en la racionali-

dad deliberativa ordinaria, el momento liberador se desvanece antes de producirse; el sujeto soberano es dejado a su suerte en las circunstancias que se le ha enseñado a comandar.

La fragilidad moral del «yo» deontológico también aparece en el nivel de los principios de primer orden. Aquí encontramos que el yo independiente, al estar esencialmente desposeído, está demasiado despojado como para ser capaz de «merecer» en el sentido habitual (Capítulo 2), ya que los postulados del mérito presuponen seres densamente constituidos, seres capaces de posesión en un sentido constitutivo, mientras que el «yo» deontológico está totalmente despojado de posesiones de este tipo. Admitiendo esta falla, Rawls encontrará derechos en las expectativas legítimas. Si somos incapaces de mérito, al menos tenemos derecho a que las instituciones cumplan las expectativas a las que dan lugar.

Pero el principio de la diferencia exige más: comienza con el pensamiento, coherente con la perspectiva deontológica, de que los bienes que tengo son sólo accidentalmente míos. Sin embargo, concluye con el supuesto de que tales bienes conforman por lo tanto un acervo común y que la sociedad tiene una demanda anterior sobre los frutos de su ejercicio. Esto le resta poder al «yo» deontológico, o le niega su independencia. O bien mis prospectos quedan librados a merced de las instituciones establecidas para «fines sociales prioritarios e independientes» (352), que pueden o no coincidir con los míos, o debo contarme como miembro de una comunidad definida en parte por tales fines, en cuyo caso dejo de estar libre de lazos constitutivos. No podemos ser personas para quienes la justicia es primaria y al mismo tiempo personas para quienes el principio de la diferencia es un principio de la justicia.

Carácter, autoconocimiento y amistad

Si la ética deontológica no logra consumar su propia promesa liberadora, tampoco tiene éxito en brindar una explicación plausible para ciertos aspectos indispensables de nuestra experiencia moral, ya que la deontología insiste en que el ser humano debe percibirse como un «yo» independiente, en el sentido de que su identidad no está nunca vinculada con sus ob-

jetivos y vínculos. Dada nuestra «fuerza moral para formar, revisar y perseguir racionalmente una concepción del bien» (Rawls 1980: 544), la continuidad de nuestra identidad está asegurada sin problemas. Ninguna transformación de mis objetivos y vínculos afectaría a la persona que soy, ya que ninguna vinculación de este tipo, por más profunda que sea, podría comprometer mi identidad.

Pero no podemos considerarnos como independientes de esta manera sin un alto costo para las lealtades y convicciones cuya fuerza moral reside en parte en el hecho de que en nuestra vida cumplirlas es inseparable de nuestra comprensión de nosotros mismos como las personas particulares que somos: como miembros de esta familia, comunidad, nación o pueblo; como sucesores de esta historia; como hijos o hijas de aquella revolución; como ciudadanos de esta república. Las lealtades como éstas son más que valores que yo pueda tener u objetivos que yo «abrazo en cualquier momento dado». Van más allá de las obligaciones a las que me comprometo voluntariamente y de los «deberes naturales» que debo a los seres humanos en tanto tales. Admiten que yo deba más a algunos de lo que la justicia exige o incluso permite, no en razón de acuerdos que yo haya establecido sino en virtud de aquellas lealtades o compromisos más o menos duraderos que, tomados en su conjunto, definen en parte a la persona que soy.

El imaginar una persona incapaz de lazos constitutivos como éstos no equivale a concebir un agente idealmente libre y racional, sino a imaginar una persona completamente falta de carácter, sin profundidad moral. Poseer carácter significa saber que me muevo en una historia que ni exige ni ordena, sino que conlleva consecuencias que no se ven afectadas por mis elecciones o mi conducta. Me acerca a algunos y me aleja de otros; vuelve más adecuados a algunos objetivos y menos a otros. En tanto ser con autocomprensión, soy capaz de reflexionar sobre mi historia y en este sentido distanciarme de ella, pero la distancia es siempre precaria y provisional, y nunca se asegura finalmente el punto de reflexión fuera de la historia misma. Una persona con carácter conoce así que está implicada de varias maneras incluso cuando reflexiona, y siente el peso moral de lo que conoce.

Esto marca una diferencia para la agencia y el autoconocimiento, ya que, como hemos visto, el «yo» deontológico, al estar completamente desprovisto de carácter, es incapaz de au-

toconocimiento en ningún sentido moralmente relevante. Cuando el «yo» no está restringido y está esencialmente desposeído, no hay sobre qué pueda cavilar una persona en la autorreflexión. Éste es el motivo por el cual, en la perspectiva deontológica, la deliberación sobre los fines sólo puede ser un ejercicio de arbitrariedad. Ante la ausencia de lazos constitutivos, la deliberación funciona como una «elección puramente preferente», lo que significa que los fines que perseguimos, estando inmersos en la contingencia, «no son relevantes desde el punto de vista moral» (Rawls 1975: 537).

Cuando actúo basado en cualidades de carácter más o menos perdurables, por el contrario, mi elección de fines no es arbitraria de la misma manera. Al consultar mis preferencias, no sólo tengo que medir su intensidad sino también evaluar su adaptación a la persona que yo (ya) soy. Me pregunto, a medida que delibero, no sólo qué es lo que realmente deseo sino también quién soy realmente, y esta última pregunta me lleva más allá de la atención de mis deseos únicamente, hacia la reflexión sobre mi identidad misma. Aunque los contornos de mi identidad estarán en cierta medida abiertos y sujetos a revisión, no estarán totalmente privados de forma. Y el hecho de que no lo estén me permite discriminar entre mis deseos y aspiraciones más inmediatos; algunos parecen ahora esenciales, otros meramente incidentales para mi definición de proyectos y compromisos. Aunque puede haber una cierta contingencia última en el hecho de que haya resultado ser la persona que soy —sólo la teología puede saberlo con seguridad— existe sin embargo una diferencia moral en que yo, en tanto la persona que soy, afirme ciertos fines en lugar de otros, tome un camino en lugar de otro. Aunque la noción de los lazos constitutivos puede parecer a primera vista un obstáculo para la agencia (el «yo», ahora restringido, ya no es estrictamente anterior), alguna fijeza relativa de carácter parece esencial para impedir la caída en la arbitrariedad que la concepción del «yo» deontológico no logra evitar.

La posibilidad del carácter en el sentido constitutivo es también indispensable para cierto tipo de amistad, la amistad marcada por percepciones mutuas además de sentimientos. En cualquier perspectiva, la amistad está vinculada con ciertos sentimientos. Nuestros amigos nos gustan; sentimos afecto por ellos, y les deseamos el bien. Esperamos que sus deseos se cumplan, que sus planes tengan éxito, y nos comprometemos de diversas maneras a procurar sus fines.

Para las personas que presumimos incapaces de lazos constitutivos, sin embargo, los actos de amistad como éstos se enfrentan con poderosas restricciones. No importa cuánto le desee el bien a un amigo, y cuán preparado esté para procurarlo; sólo el amigo por sí mismo puede saber en qué consiste ese bien. Este acceso restringido al bien de los demás se deriva del alcance limitado para la autorreflexión, que traiciona a su vez la fragilidad del «yo» deontológico. Aunque la deliberación sobre mi bien no significa más que atender a los deseos y aspiraciones dados directamente a mi conocimiento, es algo que debo hacer por mí mismo; no requiere ni admite la participación de otros. Cada acto de amistad se transforma así en parasitario de un bien identificable anticipadamente. «La benevolencia y el amor son nociones de segundo orden: buscan aumentar el bien de los individuos amados que ya está dado» (191). Incluso los sentimientos más amistosos deben esperar un momento de introspección que en sí mismo es inaccesible para la amistad. Esperar más de un amigo, u ofrecer más, sólo aparece como una presunción contra la intimidad última del autoconocimiento.

Para las personas restringidas en parte por la historia que comparten con otros, al contrario, el conocerse a sí mismas es algo más complicado. También es algo menos estrictamente privado. Aunque la búsqueda de mi bien está vinculada con la exploración de mi identidad y la interpretación de la historia de mi vida, el conocimiento que busco no es tan transparente para mí ni tan opaco para los demás. La amistad se vuelve una manera de conocer, y no sólo de querer. Cuando no estoy seguro de qué camino tomar, consulto a un amigo que me conoce bien, deliberamos juntos, ofreciendo y evaluando por turnos las diversas descripciones de la persona que soy, y de las alternativas que tengo en cuanto a su relación con mi identidad. Tomar en serio esta deliberación es admitir que mi amigo puede comprender algo que se me ha escapado, o puede ofrecer una explicación más exacta de la manera en que se compromete mi identidad en las alternativas que tengo. Adoptar esta nueva descripción es verme de una nueva manera; mi anterior imagen de mí mismo parece ahora parcial o interferida, y debo admitir en retrospectiva que mi amigo me conocía mejor de lo que me conozco a mí mismo. Deliberar con amigos es admitir esta posibilidad, que presupone a su vez un «yo» de constitución más rica de lo que la deontología admite. Aunque

por supuesto sigue habiendo momentos en que la amistad exige una deferencia hacia la imagen de sí mismo que tiene el amigo, por equivocada que esté, también esto exige conocimiento: en este caso la necesidad de ser deferente implica la habilidad de conocer.

El vernos como nos vería la deontología es privarnos de aquellas cualidades de carácter, reflexión y amistad que dependen de la posibilidad de proyectos y lazos constitutivos. Y vernos como dados a compromisos como éstos es admitir un denominador común más profundo del que describe la benevolencia, un denominador común de autocomprensión compartida a la vez que de «afectos ampliados». A medida que el «yo» independiente encuentra sus límites en aquellos objetivos y lazos de los cuales no puede apartarse, la justicia encuentra sus límites en aquellas formas de comunidad que involucran a la identidad además de los intereses de los participantes.

La deontología podría responder finalmente a todo esto con una concesión y con una distinción: una cosa es admitir que «los ciudadanos en sus asuntos personales [...] tienen lazos y amores de los que creen que no podrían, o no querrían, separarse», y que ellos «consideran impensable [...] el percibirse sin ciertos compromisos y convicciones religiosos y filosóficos» (Rawls 1980: 545). Pero en la vida pública es diferente. En este caso, ninguna lealtad u obediencia puede ser tan esencial para nuestro sentido de quiénes somos. A diferencia de nuestros vínculos con la familia y los amigos, la devoción hacia una ciudad o nación, hacia un partido o causa, no podría ser lo suficientemente profunda como para ser definitoria. Al contrario que en nuestra identidad privada, nuestra «identidad pública» como personas morales «no se ve afectada por los cambios en el tiempo» de nuestras concepciones del bien (Rawls 1980: 544-545). Aunque podemos ser seres de constitución densa en lo privado, debemos ser completamente no restringidos en lo público, y es allí donde prevalece la primacía de la justicia.

Pero una vez que recordamos la categoría especial del postulado deontológico, no resulta claro cuál podría ser el fundamento para esta distinción. Puede parecer a primera vista una distinción psicológica; el distanciamiento surge más fácilmente en la vida pública, en donde los lazos que tenemos son típicamente menos perentorios. Puedo apartarme más fácilmente, por ejemplo, de mis lealtades partisanas que de ciertas lealtades y afectos personales. Pero como hemos visto desde el

principio, el supuesto deontológico de la independencia del «yo» no puede limitarse a ser un supuesto de la psicología o de la sociología. De otra manera, la primacía de la justicia dependería del grado de benevolencia y sentimientos fraternales que una sociedad dada lograra inspirar. La independencia del «yo» no significa que yo puedo, como una cuestión psicológica, invocar en tal o cual circunstancia el distanciamiento necesario para apartarme de mis valores y fines, sino más bien que debo considerarme como el portador de un «yo» distinto de mis valores y fines, cualesquiera que sean. Es sobre todo un supuesto epistemológico, y tiene poco que ver con la intensidad relativa de los sentimientos asociados con las relaciones públicas o privadas.

Entendida como un postulado epistemológico, sin embargo, la concepción deontológica del «yo» no puede admitir la distinción requerida. Admitir posibilidades constitutivas cuando hay fines «privados» en juego parecería inevitablemente admitir al menos la posibilidad de que los fines «públicos» puedan no ser constitutivos tampoco. Una vez que los límites del «yo» ya no están fijos, individualizados por anticipado y dados con anterioridad a la experiencia, no podemos saber en principio qué tipos de experiencias podrían formarlo o deformarlo, no hay garantía de que sólo los eventos «privados» y no los «públicos» puedan concebirse como decisivos.

No son los egoístas, sino los extraños, algunas veces benévolos, los que conforman la ciudadanía de la república deontológica; la justicia encuentra su ocasión debido a que no podemos conocernos entre nosotros, o a nuestros fines, lo suficientemente bien como para gobernar por el bien común solamente. No es probable que esta condición desaparezca completamente, y en tanto no lo haga, la justicia seguirá siendo necesaria. Pero tampoco está garantizado que esto siempre sea lo predominante, y en tanto no sea así, la comunidad será posible, y constituirá una presencia perturbadora para la justicia.

El liberalismo nos enseña a respetar la distancia entre el «yo» y sus fines; cuando se pierde esta distancia, nos vemos sumergidos en circunstancias que dejan de ser nuestras. Pero al tratar de asegurar esta distancia de manera demasiado completa, debilita y socava sus propios hallazgos. Al poner al «yo» más allá del alcance de la política, hace de la agencia humana un artículo de fe en lugar de un objeto de atención y preocu-

pación continuas, una premisa de la política en lugar de su precaria conquista. Esto elude el *pathos* de la política y a su vez sus más inspiradoras posibilidades. No tiene en cuenta el peligro que plantea la política cuando funciona mal, ya que no sólo es probable que ocasione desilusiones sino también desarticulaciones. Y olvida la posibilidad de que cuando la política funciona bien, podemos conocer un bien en común que no podríamos conocer en soledad.

Bibliografía

A continuación aparecen todas las referencias indicadas en el texto, y algunas obras cuya influencia está presente en el texto aunque no se haga referencia directa a ellas.

Ackerman, B.A. 1980. *Social Justice in the Liberal State*, New Haven. [*La justicia social en el estado liberal*, Madrid, Centro de Estudios Constitucionales, 1993.]

Arendt, H. 1958. *The Human Condition*, University of Chicago Press, Chicago. [*La Condición Humana*, Barcelona, Paidós Ibérica, 1998.]

Aristóteles. *Nicomachean Ethics*. Traducción al inglés de D. Ross, 1925, Londres. [*Etica a Nicómaco* (traducción de Julián Marías y María Araujo), Madrid, Centro de Estudios Constitucionales, 1994.]

Aristóteles. *Politics*. E. Barker (comp.), 1946, Londres. [*Política*, Madrid, Centro de Estudios Constitucionales, 1970.]

Atiyah, P.S. 1979. *The Rise and Fall of Freedom of Contract*, Oxford University Press, Oxford.

Barker, E. (comp.). 1948. *Social Contract*. Nueva York.

Bartlet, V. 1915. «The Biblical and early Christian idea of property», en: C. Gore (comp.), *Property: Its Duties and Rights*, Londres.

Beer, S. 1966. «Liberalism and the national idea». *The Public Interest* (otoño) número 5, páginas 70-82.

Bell, D. 1973. *The Coming of Post-Industrial Society*, Harmondsworth. [*El advenimiento de la sociedad post-industrial*, Madrid, Alianza, 1994.]

Bickel, A.M. 1975. *The Morality of Consent*, New Haven.

Dworkin, R. 1977a. *Taking Rights Seriously*, Londres. [*Los derechos en serio*, Barcelona, Ariel, 1997.]

Dworkin, R. 1977b. «Why Bakke has no case», en: *New York Review of Books*, 10 de noviembre, páginas 11-15.

Dworkin, R. 1978. «Liberalism», en: S. Hampshire (comp.), *Public and Private Morality*, Cambridge, páginas 113-43.

Feinberg, J. 1970. *Doing and Deserving*, Princeton.
Frankfurt, H. 1971. «Freedom of the will and the concept of a person», en: *Journal of Philosophy,* número 68, páginas 5-20.
Fried, C. 1978. *Right and Wrong*, Cambridge.
Fried, C. 1981. *Contract as Promise*, Cambridge.
Hampshire, S. 1959. *Thought and Action*, Londres.
Hampshire, S. 1972. «Spinoza and the idea of freedom», en: *Freedom of Mind*, Londres, páginas. 183-209.
Hampshire, S. 1977. *Two Theories of Morality*, Oxford.
Hampshire, S. 1978. «Morality and pessimism», en: S. Hampshire (comp.), *Pubic and Private Morality*, Cambridge, páginas 1-22.
Hart, H.L.A. 1979. «Between utility and rights», en: A. Ryan (comp.), *The Idea of Freedom*, Oxford, páginas 77-98.
Hegel, G.W.F. 1807. *Phenomenology of Spirit.* Traducción al inglés de A.V. Miller, 1977. Oxford. [*Fenomenología del espíritu*, Madrid, Fondo de Cultura Económica de España, 1981.]
Hegel, G.W.F. 1821. *Philosophy of Right.* Traducción al inglés de T.M. Knox, 1952. Londres. [*Obras Completas*, Madrid, Alianza, 1987.]
Hegel, G.W.F. 1899. *The Philosophy of History.* Traducción al inglés de J. Sibree, 1956. Nueva York. [*Obras Completas*, Madrid, Alianza, 1987.]
Hume, D. 1739. *A Treatise of Human Nature.* L.A. Selby-Bigge (comp.), ²1978. Oxford. [*Tratado de la Naturaleza Humana*, Barcelona, Altaya, 1994.]
Hume, D. 1777. *An Inquiry Concerning the Principles of Morals.* Edición de 1966. Illinois, La Salle. [*Investigación sobre los principios de la moral*, Madrid, Espasa-Calpe, 1991.]
Kant, I. 1781. *Critique of Pure Reason.* Primera edición. Traducción al inglés de N. Kemp Smith, 1929. Londres.
Kant, I. 1785. *Groundwork of the Metaphysics of Morals.* Traducción al inglés de H.J. Paton, 1956. Nueva York. [*Fundamentación de la metafísica de las costumbres*, México, Porrúa. 1990.]
Kant, I. 1787. *Critique of Pure Reason.* Segunda edición. Traducción al inglés de N. Kemp Smith, 1929. Londres. [*Crítica de la Razón Pura*, México, Porrúa, 1987.]
Kant, I. 1788. *Critique of Practical Reason.* Traducción al inglés de L.W. Beck, 1956. Indianapolis. [*Crítica de la Razón Práctica*, México, Porrúa, 1990.]
Kant, I. 1793. «On the common saying: 'this may be true in theory, but it does not apply in practice'», en H. Reiss, (comp.), *Kant's Political Writings.* 1970. Cambridge, páginas 61-92.
Kant, I. 1797. *The Metaphysical Elements of Justice.* Traducción al inglés de J. Ladd, 1965. Indianapolis.
Kronman, A. 1980. «Contract law and distributive justice», en: *Yale Law Journal,* número 89, páginas 472-511.

Locke, J. 1690. *Treatise of Civil Government.* C.L. Sherman (comp.), 1937, Nueva York. [*Ensayo sobre el Gobierno Civil*, Madrid, Aguilar, 1990.]
Lyons, D. (s. f.) «Nature and soundness of the contract and coherence arguments», en: N. Daniels (comp.), *Reading Rawls*, Nueva York, páginas 141-168.
MacIntyre, A. 1967. «Egoism and altruism», en: P. Edwards (comp.), *Encyclopedia of Philosophy*, Nueva York, vol. II, páginas 462-466.
MacIntyre, A. 1981. *After Virtue*, Notre Dame. [*Tras la Virtud*, Barcelona, Grijalbo Mondadori, 1987.]
Marx, K. *Early Writings.* Traducción al inglés de R. Livingstone y G. Benton, 1975, Harmondsworth.
Mill, J.S. 1849. «On Liberty», en: *The Utilitarians,* 1973, Garden City, páginas 473-600. [*Sobre la Libertad*, Madrid, Alianza, 1997.]
Mill, J.S. 1863. «Utilitarianism», en: *The Utilitarians,* 1973, Garden City, páginas 399-472. [*El Utilitarismo*, Barcelona, Altaya, 1994.]
Murdoch, I. 1970. «The idea of perfection», en *The Sovereignty of Good*, Londres, páginas 1-45.
Nagel, T. 1973. «Rawls on justice», en: N. Daniels (comp.), *Reading Rawls*, Nueva York, páginas 1-15.
Nozick, R. 1972. «Coercion», en: P. Laslett, W.G. Runciman y Q. Skinner (comps.), *Philosophy, Politics, and Society*, 4ta serie, páginas 101-135.
Nozick, R. 1974. *Anarchy, State, and Utopia*, Nueva York. [*Anarquía, Estado y Utopía*, Buenos Aires, Fondo de Cultura Económica, 1991.]
Oakeshott, M. 1962. *Rationalism in Politics*, Londres.
Pitkin, H. 1965. «Obligation and consent - I», en: *American Political Science Review,* número 66, páginas 990-999.
Rawls, J. 1971. *A Theory of Justice*, Oxford. [*Teoría de la Justicia*, Buenos Aires, Fondo de Cultura Económica, 1993.]
Rawls, J. 1972. «Reply to Lyons and Teitelman», en: *Journal of Philosophy,* número 69, páginas 556-557.
Rawls, J. 1975. «Fairness to goodness», en: *Philosophical Review,* número 84, páginas 536-554.
Rawls, J. 1977. «The basic structure as subject», en: *American Philosophical Quarterly,* número 14, páginas 159-165.
Rawls, J. 1979. «A well-ordered society», en: P. Laslett y J. Fishkin (comps.), *Philosophy, Politics, and Society.* Oxford, 5ª serie, páginas 6-20.
Rawls, J. 1980. «Kantian constructivism in moral theory», en: *Journal of Philosophy,* número 77, páginas 515-572.
Strauss, L. 1953. *Natural Right and History*, Chicago.
Spinoza, B. 1670. *A Theologico-Political Treatise.* Traducción al inglés de R.H.M. Elwes, 1951, Nueva York. [*Tratado teológico-político*, Barcelona, Planeta Agostini, 1996.]

Spinoza, B. 1677. «The Ethics». Traducción al inglés de R.H.M. Elwes, en: *The Rationalists*, 1960, Garden City, páginas, 179-408. [*Etica*, Madrid, Aguilar, 1961.]
Taylor, C. 1975. *Hegel*, Cambridge.
Taylor, C. 1977. «What is human agency?», en: T. Mischel (comp.), *The Self: Psychological and Philosophical Issues*, Oxford, páginas 103-135.
Unger, R. 1975. *Knowledge and Politics*, Nueva York.
Wolin, S. 1960. *Politics and Vision*, Boston.

Índice de nombres y conceptos